

SERVICIUL NAȚIONAL DE EVALUARE ȘI EXAMINARE

EVALUAREA CURENTĂ ȘI EXAMENELE
GHID PENTRU PROFESORI

Coordonator: Adrian STOICA

București 2001

Contribuția autorilor:

Capitolul I: Roxana MIHAIL

Capitolul al II-lea: Roxana MIHAIL

Capitolul al III-lea: Nicoleta LIȚOIU

Capitolul al IV-lea: Adrian STOICA

Capitolul al V-lea: Marilena MÂNDRUȚ

Capitolul al VI-lea: Adrian STOICA

PREFAȚĂ

Am avut un coșmar. Visam că era vară și se apropiau examenele. Câte un specialist / inspector pentru fiecare disciplină a fost desemnat să scrie testele. Și s-au apucat de lucru cu o săptămână, cu câteva zile sau cu o zi înainte de prima probă de examen.... Mașina de scris se încăpățâna să nu bată toate caracterele, așa că unele au fost desenate cu pixul. Sau și mai simplu: testele au fost scrise de mână, mai ales baremele.

Au început examenele. „Nu înțelegem al treilea cuvânt de pe ultimul rând”, se agitău unii dintre profesori după ce au desigilat plicurile cu subiecte. „Nici o problemă. Sunăm la inspectorat.” Și de acolo mai departe. De obicei, centrul de examen primea un fax care ar fi trebuit să aducă lămuriri sau... fiecare se descurca după posibilități.

A venit rândul candidaților să arate ce știu. Emoții, căldură și speranțe. Dar și nedumeriri, confuzii. „Ce scriem la subiectul „Lupta de la Călugăreni”? Despre cauze, despre desfășurare, despre consecințe sau despre și cât de mult? Dar cum anume se punctează?”

După aceea, corectorii au trebuit să facă față greșelilor științifice și greșelilor de exprimare, omisiunilor. Dar nu numai celor din lucrările candidaților, ci și greșelilor care apăreau în baremele de corectare. Din nou nici o problemă. Se expediază fax-uri, se dau telefoane...

S-au încheiat examenele. Ce aflăm? Câți admiși, câți respinși, care este media pe țară și alte feluri de statistici... Raportul examenului? Analize și concluzii? La ce bun, oricum, ca specialiști, știm bine ce avem de făcut .

M-am trezit. Din fericire a fost doar un vis urât. Nu era decât începutul toamnei, iar examenele abia se terminaseră. Cineva a avut ideea să înceapă imediat pregătirea examenelor din anul următor, mai ales că unul dintre acestea era cu totul nou. „Nu este prea devreme? Chiar avem nevoie de atâta timp pentru a pregăti un examen? Cum, în practica internațională sunt necesari cel puțin doi ani? Dar noi avem experiență, suntem specialiști.”

Atunci s-au format grupuri de profesori pe discipline care, după serioase consultări, au stabilit *obiective de evaluare, conținuturi, matrice de specificații, tipuri de itemi*, toate acestea pentru a elabora testele pentru examen. Teste concepute altfel. Și s-au publicat *modele*. S-au făcut *pretestări* și *simulări*. Fiecare elev a primit *propriul test*, îngrijit tipărit. Elevii au fost puși mai mult în situația de a gândi și mai puțin de a reproduce simple cunoștințe. N-au mai apărut greșeli în examenele reale. Și au fost făcute *analize* și s-au publicat *rapoarte*. Cei responsabili examenele au învățat din fiecare etapă parcursă. Iar procesul a fost *transparent* pentru elevi, pentru părinți, pentru cadrele didactice, pentru publicul larg.

Cineva a avut apoi ideea să transfere o parte din experiența acumulată în pregătirea examenelor către organizarea *evaluărilor naționale* (evaluări pe eșantioane reprezentative la clasa a IV-a) și *internaționale* (PISA), dar mai ales către *evaluarea curentă*. S-au desfășurat *stagii de formare* cu profesorii, *seminarii* cu specialiștii s-au publicat *ghiduri* în sprijinul cadrelor didactice. S-a întărit colaborarea *curriculum-evaluare-formare*. A început o *reformă* reală în domeniul evaluării.

Lucrarea de față abordează toate aceste aspecte.

M-am trezit și realitatea arată altfel decât visul. Mult mai bine, cel puțin pentru folosul școlii. Dar visul nu mă dusesese cu mult timp înapoi. Doar cu câțiva ani.

Capitolul I. Evaluarea educațională și evaluarea performanțelor/rezultatelor școlare

1.1. Evaluarea educațională - de la teorie la practica educațională

Într-un sistem educațional aflat în schimbare, **evaluarea** (privită, pentru început, la modul global) și **examinarea** (desemnând atât funcția de selecție, cât și pe cea de certificare) reprezintă un adevărat loc geometric al variatelor tendințe, decizii, idei-forță, mișcări de suprafață sau de adâncime. Evaluarea în școală, secondată de examinare în diferite momente ale parcursului educațional, are o miză ale cărei costuri și al cărei impact au devenit din ce în ce mai vizibile și mai importante odată cu luarea deciziilor de politică educațională specifice reformei reale, în desfășurare în România în ultimii ani.

Documentele de politică educațională ce fac referire la domeniul evaluării și examinării din țara noastră au constituit obiectul multor dezbateri publice cu caracter general, dar și aplicat, specific, profesionist. Practica educațională însă a operat o selecție uneori diferită în ceea ce privește elementele concrete care funcționează în mod real și eficient în sistem, raportându-le pe acestea din urmă la un trecut mai mult sau mai puțin apropiat în timp, dar mai ales în concepție.

În mod evident, locul (ca și multiplele roluri ale evaluării) într-un sistem aflat în reformă este unul critic, de recunoașterea acestei realități de către toți cei implicați depinzând atât succesul, cât și autoreglarea continuă, pe viitor, a sistemului ca întreg.

Așadar, evaluarea, privită generic, îndeplinește un set de funcții și roluri, necesar a fi cunoscute și asumate, fiind în același timp: **mijloc** de relaționare eficientă a formării elevilor cu necesitățile societății aflate într-un anumit moment al evoluției sale; **modalitate esențială de control al impactului investițiilor** financiare și, eventual, de altă natură, pe care societatea le face în sistemul educațional; cel mai accesibil **mecanism de autocontrol**, realizând o cunoaștere transparentă a stării sistemului, precum și descrierea, în termeni de politică educațională, a efectului de feedback și a impactului acestuia; subsistemul **educațional** creditat cu rolul major de activare și de punere în practică a relațiilor complexe dintre directorii-manageri de școală, educatori (priviți la modul generic), elevi și părinți.

Caracteristica cea mai importantă a actualizării acestor funcții și roluri este globalitatea modului lor de angrenare în funcționare, aceasta constituind însăși natura caracterului sistemic.

Natura comprehensivă a reformei educației din România face din domeniul evaluării și examinării una dintre componentele unde schimbările de structură au devenit deja vizibile.

Obiectivele reformei în domeniul evaluării și examinării vizează:

- **evaluarea instituțională** (vizând sistemul ca întreg, precum și componentele sale — diferitele categorii de instituții de învățământ);
- **evaluarea curriculum-ului** (cu toate componentele sale);
- **evaluarea proceselor de instruire;**
- **evaluarea rezultatelor școlare** (rezultatele elevilor, din perspectivă cognitivă, afectivă și psihomotorie);
- **evaluarea personalului didactic** (din perspectiva profesionalismului — a eficienței și a relevanței acțiunii educaționale specifice întreprinse).

Procesul de evaluare, privit aici la modul global, primește astfel valențe și dimensiuni noi, pe care le putem formula astfel:

- atingerea unui echilibru dezirabil din perspectiva dezvoltării armonioase a educabililor, între cele trei domenii — cognitiv, afectiv și psihomotor;
- orientarea acțiunilor evaluative cu precădere către zonele valorilor „personalizate”, ale creativității, gândirii critice, interpretări și manifestării individuale, ale acțiunii autonome, independente etc.;
- reorientarea sistematică a actului de evaluare către competențe și capacități, abilități și atitudini în contextul unor situații și sarcini de lucru/de evaluare autentice, cât mai apropiate de viața reală,

privind viitorul educabililor și spre imaginea concretă a acestora fie în ciclul educațional următor, fie pe piața muncii.

Un element esențial al profesionalizării domeniului evaluării și examinării la noi prin reforma educației este diversificarea și specializarea **agenților evaluării** responsabili cu implementarea politicii educaționale specifice. Astfel, înființarea **Serviciului Național de Evaluare și Examinare reprezintă** pasul esențial în vederea asigurării **proiectării**, gestionării și monitorizării subsistemului evaluării și examinării, în condițiile asigurării respectării unor standarde de calitate cu adevărat europene.

În acest context, atribuțiile SNEE reprezintă dimensiunile firești ale **profesionalizării evaluării și examinării** într-un sistem educațional modern:

- a) asigură coordonarea științifică și profesională a sistemului de examene din învățământul preuniversitar, inclusiv elaborarea probelor de examen;
- b) elaborează regulamente de organizare și desfășurare a examenelor din învățământul preuniversitar, pe care le supune aprobării Ministerului Educației Naționale;
- c) efectuează anchete și sondaje de evaluare a rezultatelor școlare și a eficienței învățământului, furnizând ministerului un feedback sistematic și operativ, în scopul adoptării măsurilor de ameliorare continuă a activității de învățământ;
- d) elaborează rapoarte privind rezultate înregistrate la examenele școlare pentru informare publică și evidențiază oportunitatea adoptării de către minister a unor decizii privind factorii care permit optimizarea performanțelor școlare;
- e) organizează un sistem de evidență a rezultatelor școlare înregistrate la finalul unor cicluri de instruire, profiluri de formare sau la unele discipline școlare, pentru cunoașterea dinamicii acestora;
- f) acordă sprijin tehnic și de specialitate în realizarea unor acțiuni de evaluare întreprinse de minister, inspectorate școlare, unități de învățământ;
- g) constituie grupuri de lucru compuse din specialiști în evaluare și profesori cu experiență didactică bogată în materie de evaluare, pe bază de contract, în scopul elaborării unor bănci de itemi, pe discipline de studiu, pentru probe de evaluare a rezultatelor școlare, punându-le la dispoziția autorităților școlare și a profesorilor din rețeaua școlară;
- h) elaborează și realizează în perspectivă un program de alcătuire a unor teste standardizate pentru principalele discipline din diferite cicluri de învățământ;
- i) cooperează cu Departamentul de pregătire a personalului didactic și participă la realizarea unor programe de studii aprofundate în domeniul evaluării educaționale;
- j) organizează programe de pregătire a examinatorilor și a altor persoane implicate în desfășurarea examenelor și sprijină acțiunile de perfecționare a personalului didactic în realizarea proceselor de evaluare curentă a progresului școlar;
- k) desfășoară activități de cercetare în domeniul evaluării educaționale.

Reforma din domeniul evaluării și examinării este una dintre componentele esențiale ale reformei învățământului din România. Reforma reală nu se poate întemeia decât pe un echilibru dinamic cu schimbările introduse în celelalte componente — Curriculum, Formarea profesorilor, Managementul învățământului, Standardele ocupaționale etc. — iar miza sa specifică, cea a profesionalizării domeniului, este de natură să stimuleze acest echilibru.

În cadrul culturii educaționale, conceptul de evaluare cunoaște în societatea contemporană o multitudine de valențe, utilizări și actualizări. Funcție de con-textul în care se dorește a fi operaționalizată și de utilizatorul căruia i se adresează cu precădere, definiția este utilă în măsura în care răspunde nevoii de deschidere a câmpului cunoașterii specifice domeniului.

O definiție cu caracter „general”

„... termenul evaluare se referă la o examinare riguroasă, atentă a unui curriculum educațional, a unui program, a unei instituții, a unei variabile organizaționale sau a unei politici specifice”.

O definiție cu caracter „reglator”, unificator pentru domeniul evaluării și examinării: „Evaluarea este operația care vizează determinarea de o manieră sistematică și obiectivă a impactului, eficacității, eficienței și pertinentei activităților în relație cu obiectivele lor, pe de o parte în vederea

ameliorării activităților în curs și, pe de altă parte, în vederea planificării, programării și luării deciziilor viitoare”.

O definiție din perspectivă instrucțională: „Evaluarea este un proces sistematic de determinare a măsurii în care obiectivele instrucționale sunt atinse, realizate de către elevi”.

O definiție „pragmatică”

„Practica evaluării implică o colectare sistematică a informațiilor despre activitățile, caracteristicile și obiectivele programelor, pentru a fi utilizate de către anumiți specialiști în scopul reducerii gradului de relativitate, pentru a îmbunătăți eficiența și pentru a lua decizii cu privire la ceea ce realizează aceste programe sau la realitățile pe care le afectează. Această definiție a evaluării accentuează colectarea sistematică a informațiilor despre o gamă largă de teme pentru a fi utilizate de către anumiți specialiști într-o varietate de scopuri.

Această ultimă definiție atrage atenția asupra importanței colectării datelor care se intenționează a fi — și sunt și în realitate — folosite pentru îmbunătățirea pro-gramelor și pentru luarea deciziilor”.

Analiza definițiilor prezentate evidențiază următoarele:

1. Evaluarea educațională este o formă de evaluare întemeiată pe specificarea unui set de obiective.
2. În baza atingerii/realizării acestor obiective, părțile interesate pot face judecăți de caracter obiectiv.
3. Caracterul obiectiv al acestor judecăți se întemeiază pe interpretarea datelor reale, colectate pe parcursul procesului.
4. Judecățile de valoare elaborate pot fi folosite atât la optimizarea continuă a procesului, cât și la luarea deciziilor cu caracter de politică educațională.

Așadar, **evaluarea educațională este procesul de colectare sistematică, orientată de obiectivele definite, a datelor specifice privind evoluția și/sau performanța evidențiate în situația de evaluare, de interpretare contextuală a acestor date și de elaborare a unei judecăți de valoare cu caracter integrator care poate fi folosită în diverse moduri, prespecificate însă în momentul stabilirii scopului procesului de evaluare.**

Acest proces complex are mai multe dimensiuni sau fațete, care apar evidente în momente diferite sau la actualizarea anumitor valențe ale evaluării, astfel:

- **dimensiunea tehnică** este vizibilă în **măsurare**, ca apreciere a unei trăsături, a unei caracteristici sau a unui comportament prin raportarea la o scală de măsurare clar definită.
- **dimensiunea teleologică** este esențială pentru orice proces de evaluare, care își capătă dreptul de existență doar în momentul în care îi este formulat cu claritate **scopul**, în funcție de acesta fiind proiectate apoi obiectivele, selectate procedurile și tehnicile cele mai adecvate, construite instrumentele etc.
- **dimensiunea axiologică** este operantă în structura de adâncime a procesului: acesta este în același timp creare și atribuire de valori, contribuind atât etic, cât și deontologic la plasarea într-un sistem general uman la care cultura evaluării își aduce o contribuție specifică.

În acest context, termenul de *assessment*, provenit în engleză din latinescul *assidere* („a ședea alături de cineva spre a-l ajuta, a asista”, dar și „a ședea în instanță, ca judecător, asesor”) desemnează acțiunea de apreciere a progresului educațional/școlar, prin aprecierea unei trăsături, a unei caracteristici sau a unui comportament observabil într-o anumită situație sau într-un context precizat.

Modelul care a stat și stă în continuare la baza relaționării conceptelor menționate, precum și a structurării domeniului evaluării și examinării este **modelul tradițional al obiectivelor instrucționale**, inițiat în 1949 de către Ralph Tyler și rafinat, optimizat, adaptat continuu în mai bine de jumătate de secol de practică educațională. Acest model se organizează în jurul obiectivelor curriculare (sau al celor declarate ca aparținând unui anumit program), prespecificate, se bazează pe construirea unor instrumente de evaluare care să măsoare realizarea acestor obiective și presupune convingerea că parcursul „obiective instituite — experiențe de învățare — evaluare” este unul coerent și suficient sieși.

De-a lungul timpului, modelele ce stau la baza construirii unui proces de evalua-re în sens foarte larg — de la evaluarea de proces la evaluarea programelor educaționale și/sau a implementării acestora — s-au polarizat astfel:

A. Modele tradiționale, „convenționale”:

- **Modelul centrat pe obiective** (Tyler - Mager)

- **Modelul centrat pe decizie** (CIPP = Context, Intrare, Proces, Produs), unde funcția de decizie o are inițiatorul sau beneficiarul procesului.

B. Modele complementare, „nonconvenționale”

- **Modelul evaluării fără obiective** (care recuperează în favoarea celui evaluat efectele neintenționate ale situației de învățare sau chiar ale situației de evaluare, complementând astfel limitările posibile induse de formularea inițială a obiectivelor).

- **Modelul evaluării responsive** (care centrează procesul de evaluare pe realitatea din clasă, bazată pe interacțiune și negociere, punându-l pe baze preponderent calitative).

- **Modelul evaluării iluminative** (care întemeiază procesul de evaluare pe caracteristici exclusiv calitative, precum: naturalistică, euristică, interpretativă, individualizatoare etc.).

- **Modelul evaluării naturalistice** (care oferă o abordare holistică a obiectivelor și subproceselor programului de evaluare și accentuează, explicându-le, efectele și semnificațiile procesului).

- **Modelul evaluării calitative** (care dezvoltă o cunoaștere reflexivă privind contextele situației educaționale, urmărind parcursul în patru etape: observație, descriere, interpretare, estimare — engl. *appraisal*).

Fiecare dintre modelele trecute în revistă vizează un anumit scop sau asocieri de scopuri ce conferă caracteristicile dominante ale procesului de evaluare. Succesul procesului este dat de gradul de coerență atins între: scopuri, obiective, instrumente de evaluare; de adecvarea demersurilor proiectate; de consistența „proiectului de evaluare”, de claritatea comunicării către toți cei interesați a interpretărilor și concluziilor finale.

În proiectarea oricărui proces de evaluare stabilirea scopului constituie pasul esențial. Macintosh formulează șase scopuri posibile pentru un proces de evaluare educațională, astfel:

„(a) **Diagnostic** — pentru a monitoriza progresul și pentru a afla cum anume asimilează elevul ceea ce i-a fost predat. Ca un rezultat al evaluării diagnostice pot fi întreprinse acțiuni specifice.

(b) **Evaluativ** — pentru a evalua eficiența predării, ceea ce, din nou, poate conduce la o acțiune specifică.

(c) **Orientare/consiliere** — pentru a asista elevii în luarea deciziilor privind viitorul, fie că acestea privesc alegerea unei discipline sau a unui anumit curs, fie că ajută la alegerea unei cariere potrivite.

(d) **Predictiv** — pentru a descoperi abilități și aptitudini potențiale și pentru a prevedea succesele viitoare probabile, fie în cadrul școlii, fie în afara ei.

(e) **Selecție** — pentru a determina care sunt cei mai potriviți candidați pentru un anumit curs, o anumită clasă, sau pentru studiile universitare.

(f) **Notare** — pentru a repartiza elevii în grupuri, pentru a discrimina între indivizii din cadrul aceleiași grup.”

Opțiunea pentru un anumit scop în momentul proiectării demersului evaluativ și al planificării și structurării procesului atrage după sine necesitatea clarificării modului de raportare/interpretare a datelor privind performanța vizată. Cele două moduri principale se situează într-o dihotomie operațională:

Modul normativ

- presupune realizarea de comparații între indivizi privind performanța /performanțele;
- implică proiectarea unor instrumente de evaluare care să raporteze **performanța relativă** a unui elev în raport cu restul populației;
- nu va spune ceea ce un anumit individ sau un elev poate face/realiza.

Modul criterial

- presupune realizarea de comparații în baza unor criterii externe, prestabilite;

- implică proiectarea unor instrumente de evaluare care să raporteze **performanța absolută** a unui elev, în termenii a ceea ce acesta poate face;
- nu va putea plasa elevii **în ordinea** dată de gradul de abilitate evidențiat. d

Dar practica educațională nu operează la modul ideal cu dihotomii operaționale teoretice, ci le actualizează contextual: evaluarea implică atât **elemente normative**, cât și **elemente criteriale**. În plus, atât instrumentele normative, cât și cele criteriale trebuie să fie proiectate în baza unor obiective de evaluare. De asemenea, performanța elevului poate fi judecată și în funcție de rezultatele anterioare, urmărindu-se evoluția în timp. În fine, interpretarea rezultatelor obținute în urma aplicării unui instrument de evaluare proiectat conform specificațiilor de tip preponderent **normativ** sau **criterial** va fi realizată în consecință. De exemplu, un test normativ va acoperi o zonă curriculară mai extinsă, în vreme ce un test criterial va viza mai puține concepte, dar o va face în adâncime. **Un test normativ** va fi preferabil într-o situație de selecție a candidaților **pentru** un număr limitat de locuri, în vreme ce **un test criterial** va fi necesar într-o situație în care este vizat nivelul minim de performanță atins de un număr cât mai mare de elevi. O evaluare cu caracter normativ va distribui subiecții de-a lungul întregii scale de corectare/notare, în vreme ce o evaluare criterială va oferi o măsurare a nivelului de stăpânire atins de grup în raport cu un criteriu clar specificat.

Întrebări clasice — cum trebuie să se proiecteze evaluarea în termeni de **moduri** adecvate scopului — i-a răspuns în 1977 Rowntree, (citată de Macintosh, op. cit., p. 16-17). Modurile descrise funcționează de cele mai multe ori sub forma unor dihotomii operaționale, dar acestea nu sunt nici absolute și nici exclusive, căci caracteristicile intrinseci ale modurilor se suprapun uneori în actul real de evaluare. Cele opt dihotomii uzuale sunt următoarele:

- formal vs. informal • proces vs. produs
- formativ vs. sumativ • intern vs. extern
- continuu vs. final • convergent vs. divergent
- evaluarea activităților curente vs. examinare • idiografic vs. nomotetic

Așadar, **adecvarea la scop** (engl. *fitness for purpose*) implică alegerea **modului** cel mai potrivit, ansamblu de decizii care se iau în funcție de următoarele criterii:

- scopul major al evaluării/aprecierii/măsurării educaționale;
- constrângerile și determinările logistice (aici incluzându-se și costurile umane și materiale implicate);
- natura și dimensiunile grupului-țintă investigat;
- scopurile, natura și caracteristicile curriculum-ului sau programului ce întemeiază procesul de evaluare/examinare;
- importanța, pentru toți cei implicați, a mizei rezultatelor așteptate;
- forma de prezentare, în funcție de audiența sau de beneficiarii raportului final.

Abia după luarea deciziilor privind aceste probleme se poate trece la selectarea tehnicilor de evaluare ce urmează a fi folosite (de exemplu, a formatului de itemi de utilizat pentru un anumit test).

Ca proces, **evaluarea educațională** poate fi abordată din două perspective. Macintosh (op. cit., p. 10) le descrie ca având următoarele caracteristici:

„Abordarea pragmatică

- Evaluarea a ceea ce se „întâmplă” de fapt în clasă.
- Analiza rezultatelor evaluării pentru a discrimina între elevi și pentru a asigura o evaluare bine echilibrată.
- Adaptarea tehnicilor evaluării pentru a face față ocaziilor, provocărilor oferite de rezultatele neașteptate.
- Amânarea notării finale până când toate rezultatele evaluării pot fi echilibrate și ajustate în mod adecvat.

Abordarea predeterminată

- Stabilirea de finalități satisfăcătoare.

- Formularea obiectivelor prin care pot fi atinse, realizate finalitățile.
- Determinarea criteriilor cu ajutorul cărora poate fi măsurat progresul elevilor sau nivelul/gradul de stăpânire a obiectivelor.
- Pre-testarea și post-testarea materialului de evaluat. Mai întâi se stabilește dacă acesta este potrivit pentru vârsta și pentru gama de abilități ale elevilor. Apoi se asigură dacă materialul de evaluat este relevant pentru disciplina predată. În final, se asigură ca rezultatele obținute să fie luate în considerare.
- Determinarea nivelurilor finale de stăpânire a criteriilor precizate, i.e., a notelor/calificativelor de trecere."

Studiile și cercetările din domeniu semnalează faptul că cel de-al doilea tip de abordare este cel mai frecvent mai ales în cazul evaluărilor pe scară largă, precum și al examenelor naționale. Aceste abordări nu reprezintă „rețete” care trebuie să fie urmate, ci o succesiune logică de etape, necesar a fi proiectate și parcurse.

1.2. Evaluarea de proces vs. evaluarea de sistem

Pe fundalul tendinței generale de transferare a responsabilităților (și de ridicare a gradului de conștientizare instituțională) într-o măsură din ce în ce mai semnificativă direct către autoritățile educaționale locale, ca și instituțiilor școlare, problema **eficienței educaționale devine din ce în ce mai mult o constantă pe agenda** factorilor de decizie ce conturează politica educațională în diferite sisteme educaționale, punându-se din ce în ce mai acut în termenii eforturilor de transparentizare a costurilor educației, de comunicare specifică a acestora atât factorilor decizionali, cât și opiniei publice, de activare a unor mecanisme de complementare a eforturilor bugetare prin surse de autofinanțare, de gândire a strategiilor coerente pe termen lung privind relațiile educației cu celelalte subsisteme sociale.

Conceptul de **eficiență educațională face legătura** dintre evaluarea de proces și evaluare de sistem, făcând vizibile și explicite relațiile dintre cele două. Astfel, „eficiența educațională are o sferă mai largă decât aprecierea școlară (docimologică) și decât eficiența economică, pe care o integrează într-un sistem de corelații cu multiple ramificări”. Aprecierea eficienței educaționale în termeni cantitativi sau calitativi, în funcție de demersul ales, se realizează la nivelul „macro”, al sistemului de învățământ, iar beneficiarul direct este societatea în ansamblu.

În dihotomia **evaluare de proces/evaluare de sistem**, cea dintâi **generează** un set de date, rezultate, elemente de control care pot fi utilizate de cea de-a doua ca indicatori pentru aprecierea **eficienței educaționale**, a sistemului ca întreg sau doar a anumitor componente ale sale. Acest lucru face ca **eficiența educațională să constituie un domeniu de cercetare și de investigație în sine**: „... un domeniu de evaluare prin analiza multivariată a rezultatelor interne și externe, cantitative și calitative ale sistemului și proceselor educaționale, precum și a determinantilor care operează pe rezultate” (op. cit., p: 20). În acest context se activează valența de monitorizare și conducere pe care o are evaluarea educațională în raport cu sistemul, facilitând, stimulând și, în ultimă instanță, determinând deciziile de reorganizare, optimizare, inovație la nivelul sistemului, ca și la cel al procesului educațional.

Aceste valențe operează sub forma cunoscutelor **funcții ale evaluării și examinării**, prezentate, în continuare, sub forma unui inventar care încearcă să sur-prindă varietatea și multitudinea posibilităților lor de utilizare în diferite situații.

Natura funcției	Sfera de operare a funcției	Instrumentele cele mai frecvent utilizate
Diagnostică	Identificarea nivelului performanței, a punctelor tari și slabe, pe domenii ale performanței	Instrumente de evaluare diagnostică: teste psihologice, de inteligență, teste de cunoștințe sau de randament etc.
Prognostică	Estimarea domeniilor sau a zonelor cu performanțe viitoare maxime ale educabililor	Teste de aptitudini, de capacități sau abilități specifice
de Selecție	Clasificarea candidaților în ordinea descrescătoare a nivelului de performanță	Ideală este utilizarea de teste standardizate de tip normativ.

	atins, într-o situație de examen sau de concurs. Funcția se poate actualiza, cu o miză mai mică, în situația necesității creării claselor de nivel.	Funcția este activată și de către anumite componente ale examenelor naționale.
de Certificare	Recunoașterea statutului dobândit de către candidat în urma susținerii unui examen sau a unei evaluări cu caracter normativ.	Eliberarea de certificate, diplome, acte dovedind dobândirea unor credite etc.
Motivațională	Activează și stimulează autocunoașterea, autoaprecierea, valențele metacognitive în raport cu obiectivele procesului educațional stabilite de la început sau în funcție de obiectivele de evaluare comunicate anterior.	<i>Feedback</i> structurat din partea profesorului-evaluator, informal (oral) sau formal (sub forma rapoartelor-comentarii)
de Consiliere	Orientează decizia elevilor și a părinților, în funcție de nivelul performanțelor obținute, astfel încât orientarea școlară și/sau profesională a elevilor să fie optimă, în echilibru stimulatîv între dorințe și posibilități.	Discuții individuale, „seri ale părinților”, vizite cu scop de familiarizare a unor instituții educaționale, alte forme de consiliere destinate elevilor sau părinților.

Tabelul nr. 1 — *Funcțiile evaluării și examinării*

În momentul în care cel care proiectează o evaluare educațională ia decizia să realizeze unul dintre tipurile de evaluare care actualizează una sau mai multe funcții, se află în situația de a răspunde unor întrebări „clasice”, de al căror răspuns depind structura demersului de evaluare, ca și eficiența actului de evaluare în sine. Succesiunea acestor întrebări poate fi următoarea:

Ce evaluăm?

Sistemul ca întreg sau doar una dintre componentele sale, randamentul sau eficiența educațională, procesul de formare/educațional, rezultatele școlare, niveluri de performanță, competențe în acțiune, aptitudini, abilități, capacități, elemente de competență, constructe mentale, reprezentări mentale, atitudini etc.

Cu ce scop evaluăm?

Formativ/sumativ, de plasament, de orientare/consiliere etc.

Pe cine evaluăm?

Educabilii (elevii, formabilii, un anumit grup de vârstă sau de abilități/competențe) etc.

Cum evaluăm?

Prin stabilirea clară și transparentă a parcursului: scopuri — obiective —instrumente de evaluare — rezultate — interpretare — comunicare.

Când evaluăm?

La începutul unui proces (ciclu educațional, an școlar, semestru, oră de curs), pe parcursul acestuia, la finalul său, după un anumit timp de la finalizarea sa.

Cu ce evaluăm?

Cu instrumente de evaluare orală/scrisă/practică

Prin observație directă, (semi)structurată, pe parcursul procesului

Prin exerciții, probleme, eseuri, teme pentru acasă

Prin proiecte, referate, teme pentru investigațiile individuale sau de grup

Prin portofolii individuale sau instituționale

Prin proceduri de autoevaluare, evaluare pe perechi și de grup, cu scopul creșterii reflecției metacognitive și al socializării instituționale

Cine beneficiază de rezultatele evaluării?

Elevii, absolvenții, profesorii-evaluatori, părinții, concepatorii de curriculum, autoritățile abilitate în proiectarea examenelor, factorii de decizie, patronatul etc.

Odată stabilite răspunsurile specifice la aceste întrebări, se poate proiecta și construi situația de evaluare sau de examinare, care capătă astfel conturul unui **proiect educațional real**.

1.3. Relația evaluare curentă — examene

Diferite sisteme educaționale rezolvă în mod diferit problema extrem de delicată a echilibrului ideal dintre **evaluarea curentă** și **examene**. În contextul actual în care, cel puțin în Europa se fac eforturi susținute pentru recunoașterea reciprocă a certificatelor și creditelor naționale, eforturile de „transparentizare” a procedurilor de evaluare în situațiile de examen au devenit o preocupare de prim ordin a factorilor de decizie.

În sistemele educaționale din societățile ce situează competiția pe primul plan al dezvoltării individuale **examenele** constituie încă „ritualuri de trecere” (*„les rites de passage”* despre care vorbesc sociologii educației) cu o putere și o importanță remarcabile. De aceea, modul în care sunt proiectate, administrate, interpretate și optimizate continuu reprezintă o adevărată cheie de boltă a sistemului educațional. Spre exemplu, la noi, în reprezentarea mentală colectivă, Bacalaureatul are încă valoarea ritualului de trecere major spre vârsta maturității ceea ce face ca proiectarea sa, în viitor, să necesite asigurarea unui demers ferm de armonizare a scopurilor, obiectivelor, tehnicilor, instrumentelor și efectelor dorite (ca și ale celor neașteptate!) ale mizei sale educaționale și sociale.

I. T. Radu oferă o definiție operațională a examenului: „Examenul este forma instituționalizată ca organizare și ca sistem de tehnici prin care progresele înregistrate de elevi sunt verificate periodic”.

Preluând observațiile lui Péron, din *Examens et docimologie*, I. T. Radu descrie funcțiile examinării astfel:

- diagnosticarea procesului;
- stimularea subiectului;
- formularea unor predicții.

În continuare, într-o analiză a examenelor din sistemul nostru educațional, li se aduc următoarele reproșuri (I. T. Radu, op. cit., p. 214):

- au un caracter de activitate de sine stătătoare, ruptă de restul proceselor de instruire;
- se face un apel excesiv, chiar exclusiv, la volumul cunoștințelor acumulate;
- este determinantă învățarea bazată pe memorie și pe tehnici automatizate;
- au o slabă valoare prognostică.

Dintre elementele pozitive ale examenelor, sunt menționate:

- acțiunea lor ca factor extern stimulator;
- faptul că uneori constituie un exercițiu de muncă util;
- faptul că oferă ocazia de a face sinteze asupra cunoștințelor acumulate;
- faptul că oferă informații privind starea educației, ca și eficiența ei.

Așadar, **avantajele** examenelor publice pot fi sumarizate astfel:

1. pot oferi o apreciere obiectivă a performanțelor educabililor (elevi sau absolvenți ai unui anumit parcurs educațional);
2. pot contribui, în timp, într-un mod specific, la definirea și monitorizarea standardelor educaționale naționale;
3. au un statut „preferențial” în ochii opiniei publice, deoarece (re)prezintă cele mai vizibile „ieșiri” ale sistemului educațional.

Punctele slabe ale examenelor publice, pot fi sumarizate astfel:

1. pot acoperi doar o parte limitată a curriculum-ului parcurs prin demersul educațional;
2. momentul examinării poate evidenția doar o parte restrânsă a abilităților, elementelor de competență și cunoștințelor celor examinați, în funcție de subiectul sau sarcina de lucru stabilite;
3. pot avea efecte negative asupra candidaților care fie nu au pregătirea psihologică necesară susținerii unui examen, fie nu performează bine în condițiile situației de examinare;

4. pot încuraja doar acele practici educaționale (instrucționale, metodice, didactice sau general pedagogice) care conduc la obținerea unor performanțe ridicate în condiții de examen;

5. pot încuraja memorarea pe dinafară a anumitor elemente despre care se știe că, „prin tradiție”, apar în examinare;

6. pot avea costuri ascunse extrem de ridicate atunci când întreprind eforturi susținute pentru a crește calitatea și profesionalismul examenului;

7. mizele prea ridicate ale anumitor examene publice pot demotiva segmente importante ale populației școlare care nu se mai prezintă la susținerea acestora

În articolul dedicat „examenelor externe” din *The International Encyclopedia of Educational Evaluation*, H. C. Macintosh identifică trei caracteristici majore ale examenelor externe (definite ca examenele care sunt proiectate, se desfășoară și generează rezultate în afara procesului propriu-zis de formare):

- extinderea substanțială a scalei lor de operare (tendința generală fiind astăzi în lume ca examenele externe să fie produse integral de agenții/instituții specializate, în condițiile respectării unor standarde de înalt profesionalism, ceea ce face posibilă comparabilitatea între examene din sisteme diferite în baza unui „cod de practică profesională” comun și întemeiază sistemul de recunoaștere reciprocă a acestora);

- accentul susținut pus pe utilizarea rezultatelor examenelor externe preponderent în scopul selecției (deși sunt recunoscute și celelalte scopuri posibile ale examinării — certificare, diagnoză, evaluare de sistem, orientare, notare, predicție — opțiunea din ce în ce mai vizibilă astăzi este pentru selecție);

- gama inevitabil relativ limitată a ceea ce este în mod real evaluat în mod concret în situația examenului extern (examelele cu miză ridicată sunt și în prezent cantonate în zona „academicului”, examenele vizând în special competențe și abilități directe legate de disciplinele tradițional academice; în plus, aceasta se asociază și cu gama restrânsă a tehnicilor de testare posibil a fi utilizate în situația de examinare).

Aceste caracteristici determină autorul citat să creeze o definiție operațională pentru „examelele externe”:

„O mare majoritate a examenelor externe actuale sunt teste normative, terminale / finale, scrise, destinate celor cuprinși în învățământul de zi, cursuri integrale. Tind să nu ia în calcul activitățile/lucrurile în grup și nici curriculum-ul interdisciplinar și sunt în primul rând preocupate să măsoare cunoștințele și/sau aptitudinile ce sunt considerate relevante pentru studiile academice ulterioare” (Walberg, op. cit., p. 503).

Reproșurile în special de factură „tehnică” ce se aduc în mod curent examenelor externe așa cum se desfășoară ele în Europa sfârșitului de mileniu vizează următoarele:

- datorită caracterului lor preponderent normativ (în care compararea performanței / performanțelor individuale se face cu cea a grupului care susține același examen), cerința principală este ca instrumentele proiectate (în speță, testele scrise) să discrimineze cât mai puternic, în defavoarea caracterului criterial pe care l-ar putea avea;

- datorită gamei restrânse a tehnicilor de testare și a instrumentelor proiectate (și acceptate „prin tradiție”), accentul cade, în mod inevitabil, pe accentuarea testării cunoștințelor, în defavoarea conceptelor, abilităților, caracteristicilor și atitudinilor personale;

- datorită mizei pe care o dețin, prin definiție, examenele pot răspunde cu greutate sau chiar deloc necesității de schimbare și nu încurajează inovația sau introducerea ei pe această cale.

La noi, în mod tradițional, **examenul** face referire la controlul rezultatelor școlare în diferite momente ale parcursului educațional, iar **concursul** actualizează funcția de selecție într-o situație specifică de evaluare, în condițiile limitărilor impuse de un *numerus clausus* rigid.

Situația de examinare se caracterizează printr-o miză socială extrem de ridicată, mai ales în condițiile **examenelor naționale**. Analizată până acum, mai ales din punctul de vedere al candidaților care se confruntă cu examenul, această miză este majoră și din punctul de vedere al societății, care investește în examenele naționale. Din perspectiva proiectării acestora, caracteristica de național se poate asigura doar prin respectarea de către toți candidații a acelorași condiții (începând cu abordarea probelor de examen, continuând cu unificarea procesului de corectare și notare și

finalizând cu asigurarea pentru toți cei implicați, a acelorași condiții de desfășurare și de raportare/comunicare a rezultatelor examenului). **Examenele locale** reprezintă opțiunea unor autorități educaționale locale, oglindind necesități și decizii în consecință. În sistemul educațional românesc, cele două **examene naționale**, **Capacitatea** și **Bacalaureatul**, sunt secondate de examene cu caracter local, precum: **admiterea în liceu**, **admiterea în învățământul superior precum și — în învățământul terțiar și cel superior, de examenele de absolvire a școlilor profesionale, postliceale, de examenele de licență** sau de cele din **sistemul studiilor aprofundate**.

Tocmai datorită mizei lor extrem de ridicate, efectele examenelor trebuie atenuate, ponderate, „modelate pozitiv” la nivelul sistemului printr-o altă modalitate de evaluare care să poată îndeplini funcțiile de control și de monitorizare a procesului educațional. De aceea, o tendință din ce în ce mai pregnantă în ultimul deceniu este de a acorda o importanță crescândă **evaluărilor naționale** (engl. *national assessments*).

Evaluările naționale au drept principală funcție **monitorizarea standardelor naționale** de performanță. Ca forme de evaluare, vizează categorii diferite de populații-țintă, furnizând date empirice ce pot fi folosite la evaluarea de sistem, precum și pentru luarea deciziilor de optimizare a procesului educațional. În *design-ul evaluărilor naționale* predomină caracterul/scopul **criterial**, căci acestea descriu ceea ce elevii **pot face**, în funcție directă de obiective, în calitatea lor de criterii stabilite la nivel național. Din rațiuni în principal financiare, evaluările naționale se fac, în cele mai multe țări, pe bază de eșantion reprezentativ și nu vizează toată populația școlară dintr-o anumită cohortă.

Astfel, în Statele Unite ale Americii, programul evaluărilor naționale are deja o tradiție reductibilă: inițiat în 1969, *National Assessment of Educational Progress* (NAEP), ca proiect mandatat de Congresul S.U.A., vizează progresul academic al elevilor, pe diferite niveluri de performanță și în diferite arii ale disciplinelor de studiu (cele „tradiționale” în sistemul educațional american: citire/lectură, matematici, știință, scriere, istorie/geografie). Vârstele la care se realizează testarea sunt: 9, 13 și 17 ani. Eșantionul național reprezentativ este de aproximativ 30.000 de elevi, iar datele sunt culese anual. Scopul principal este monitorizarea **nivelurilor performanței academice de-a lungul timpului**, accentuând în rapoartele publice tendințele constatate. Consiliul de conducere a evaluării naționale (engl. *National Assessment Governing Board*), stabilit în 1998 de către Congres, stabilește următoarele:

- „formularea instrucțiunilor oficiale de politică pentru NAEP;
- selectarea ariilor pe discipline de studiu ce vor fi evaluate, dintre cele incluse în Scopurile Educației Naționale;
- stabilirea nivelurilor adecvate de performanță ale elevilor;
- proiectarea obiectivelor de evaluare și a specificațiilor testelor printr-o abordare de tip „obținerea consensului general”;
- proiectarea metodologiei de evaluare;
- stabilirea instrucțiunilor pentru raportarea și diseminarea rezultatelor NAEP;
- proiectarea standardelor și procedurilor pentru comparațiile între state, regionale și naționale;
- determinarea adecvării itemilor testelor și garantarea lipsei caracterului lor tendențios (cultural, religios, sexist etc.)
- întreprinderea de acțiuni pentru optimizarea formei și utilizării evaluărilor naționale”.

În afara testelor propriu-zise, NAEP colectează, prelucrează și integrează și in-formații privind variabile relevante, considerate a le descrie cât mai corect și adecvat contextului sociocultural al elevilor. Astfel, „rapoartele elevilor privind experiențele lor de la școală și de acasă legate de învățarea în cazul diferitelor arii ale disciplinelor, oferă un context important pentru înțelegerea tendințelor progresului academic de-a lungul timpului” (op. cit., p. 22).

Informațiile specifice privind tendințele în evoluția nivelurilor de performanță academică a elevilor datează din 1977 la științe, din 1978 la matematică, din 1971 la citire/lectură, din 1984 la scriere. Programul a fost introdus gradat și, de-a lungul timpului au existat schimbări și modificări care însă n-au alterat scopul principal, cel de monitorizare a standardelor de performanță.

Există și sisteme educaționale în care întreaga populație școlară participă la acest proces de evaluare, cum este cazul Marii Britanii. Învățământul obligatoriu, de la 5 la 16 ani, este împărțit în patru stadii cheie (engl. *key stages*). La finalul fiecăruia dintre acestea, (KS1, 7-8 ani, KS2, 11-12 ani și KS3, 14-15 ani), toți elevii participă la o testare națională realizată în baza sarcinilor de lucru standardizate (engl. *Standard Assessment Tasks - SAT*) care sunt proiectate și publicate de către serviciile de examinare in-dependente și moderate extern (deci nu de către profesorii școlii în care se realizează testarea). *SAT* nu sunt teste standardizate, „clasice”, ci sarcini de lucru strâns legate de activitățile obișnuite desfășurate în clasă. Progresul elevului este evaluat în termenii nivelurilor precizate în curriculum-ul național. Rezultatele evaluării curente, realizate de către profesor la clasă (engl. *coursework assessment*) sunt comparate cu rezultatele obținute la *SAT* în evaluările naționale, iar **judecata** de evaluare finală a performanței elevului se finalizează în cadrul procesului de moderare externă.

În general, proiectarea unei evaluări naționale ia în calcul următoarele categorii de probleme:

- identificarea acelor aspecte ale disciplinelor de studiu care sunt bine sau slab realizate în procesul educațional;
- identificarea modului în care variază în timp nivelurile de realizare / atingere a standardelor;
- selectarea anumitor domenii ale curriculum-ului, socotite importante, și urmărirea lor în profunzime, pe parcursul unei perioade mai lungi de timp;
- definirea cu atenție a populației-țintă investigate (spre exemplu, în SUA *National Assessment of Educational Progress* este periodică, investighează anumite domenii ale curriculum-ului și este realizată pe baza unui eșantion național pentru elevii de 14 ani), corelarea cu domeniile curriculare investigate, cu abilitățile urmărite, cu modul de raportare a rezultatelor către opinia publică și cu tipul de decizii posibil a fi luate.

Categoriile de beneficiari care sunt interesați de rezultatele **evaluărilor naționale** sunt: factorii de decizie, concepatorii de curriculum, formatorii de formatori, autorii de manuale și de auxiliare școlare, specialiștii în evaluare și cei care proiectează examenele, părinții, patronatul, opinia publică și, desigur, nu în cele din urmă, presa.

În ultimii ani Serviciul Național de Evaluare și Examinare a elaborat și pus în practică un program de **evaluare națională** pe termen lung. Etapele de proiectare și implementare ale acestuia sunt următoarele:

1. Decizia asupra politicii de evaluare națională (inclusiv în privința alocării resurselor financiare necesare).
2. Decizia privind nivelurile educaționale cheie (clasele) și disciplinele de studiu care vor fi evaluate. Decizia privind datele complementare (mediu social și familia etc.) ce vor fi colectate.
3. Proiectarea și construirea cadrului de evaluare, a specificațiilor de test, a procedurilor de corectare și notare, a itemilor, a chestionarelor etc.
4. Înființarea unei baze de date privind administrarea și analiza statistică a datelor obținute.
5. Stabilirea procedurilor administrative necesare, incluzând manualele de test, formarea evaluatorilor/corectorilor etc.
6. Aplicarea de teste pilot în școlile făcând parte din eșantionul restrâns stabilit. Analizarea datelor și a procedurilor aplicate. Revizuirea instrumentelor și a procedurilor, în urma aplicării.
7. Selectarea școlilor ce constituie eșantionul reprezentativ real. Finalizarea procedurilor administrative.
8. Conducerea evaluării propriu-zise și monitorizarea „calității” acesteia.
9. Corectarea lucrărilor, introducerea datelor în baza de date, „curățarea” datelor, analizarea rezultatelor. Monitorizarea calității.
10. Scrierea rapoartelor pentru diferite tipuri de beneficiari: factori de decizie, concepatori de curriculum, profesori, opinia publică etc.
11. Diseminarea principalelor rezultate prin intermediul sesiunilor de formare etc.
12. Acționarea conform deciziilor luate.

În ultimii ani, în România au fost întreprinse trei **evaluări naționale** la sfârșitul învățământului primar, vizând nivelul anumitor performanțe specifice ale absolvenților ciclului primar. Astfel, în 1995, nivelul cunoștințelor dobândite la matematică și la limba română a fost investigat cu ajutorul unor probe de cunoștințe proiectate pe baza manualelor în uz. În 1996, anumite abilități funcționale și capacități constituind finalități ale învățământului primar au fost investigate cu ajutorul unor probe vizând gradul de alfabetizare funcțională. În 1998 a fost urmat același model ca în 1996, cu diferența că la limba română a fost investigat în mod specific **gradul** alfabetizării funcționale, iar la matematică a fost proiectată o probă de cunoștințe aplicate în situații concrete. Fiecare dintre aceste evaluări naționale a fost finalizată cu rapoarte tehnice care au fost puse la dispoziția factorilor de decizie.

Evaluarea națională realizată în 1998 a avut ca obiective pe termen lung: stabilirea mecanismelor de bază de monitorizare a standardelor naționale; delimitarea, în linii mari, a domeniilor de investigare a progresului educațional (Am termeni de capacități de bază, conținuturi asociate etc.). Pentru stadiul 1998, obiectivele au vizat, pe de o parte, obținerea de date concrete privind abilitățile și cunoștințele absolvenților clasei a IV-a la limba maternă (limba română și limba maghiară) și matematică și, pe de altă parte, argumentarea, prin date empirice, în baza performanțelor demonstrate de către subiecți, a relațiilor dintre obiectivele evaluate și unele caracteristici ale mediului educațional (privind învățătorul și școala), la finalul primului ciclu educațional.

Momentul evaluării naționale 1998 are cel puțin trei implicații majore pentru programul de evaluare națională pe termen lung, care intră într-o nouă fază odată cu proiectarea și administrarea, de către SNEE, a evaluării naționale de la finalul clasei a IV-a, din decembrie 2000:

1. Au fost furnizate acum reperele pentru definirea domeniilor și a capacităților ce vor constitui baza pentru investigațiile viitoare;

2. Este posibilă acum definitivarea formatului testelor pentru **Programul**

evaluărilor naționale pentru ciclul primar, funcție de scopul specific și

de obiectivele corelate stabilite;

3. Este posibilă acum, din perspectiva evaluării longitudinale corelarea, în bază de capacități și de elemente de competență, a curriculum-ului vechi cu cel nou introdus.

În continuare, SNEE proiectează și va implementa un program pe termen lung privind **evaluările naționale**, explorând posibilitățile de extindere a investigației și la nivelul altor clase (de exemplu, clasele a VII-a sau a VIII-a).

În contextul în care evaluărilor **naționale** le va reveni din ce în ce mai mult sarcina monitorizării standardelor de performanță, scopurile principale, dominate ale **examenelor naționale din România** vor continua să fie:

1. Certificarea performanțelor demonstrate în situația de examinare, a anumitor elemente de competență, a unor competențe, abilități sau cunoștințe;

2. Selecția — sau controlul accesului, în cazul existenței unui *numerus clausus*. Miza examenului crește invers proporțional cu numărul de locuri disputate.

Scopurile de diagnostic și de **prognoză** vor fi îndeplinite doar parțial de examenele naționale, acestea fiind preluate în timp, din ce în ce mai mult, de către evaluările naționale.

Una dintre modalitățile cele mai puternice de profesionalizare a **examenelor** naționale este proiectarea de programe de examen care reprezintă veritabile documente de instrumentare a rolului și statutului evaluării și examinării într-un sistem educațional. Dacă programele școlare, în calitatea lor de componente curriculare, stabilesc „conținuturile” procesului educațional — în mod tradițional la noi, ale învățării pe discipline —, programele de examen (engl. *examination syllabi*) trebuie să focalizeze atenția asupra a ceea ce va constitui examenul ca eveniment de evaluare, organizațional și administrativ, „consacrând” o situație de fapt în ceea ce privește performanțele elevilor/absolvenților într-un anumit moment al evoluției lor. În consecință, programele de examen sunt documente complexe, care trebuie să fie cunoscute nu numai de către profesori, ci și de către elevi, cu cel puțin 1-2 ani înainte de „evenimentul — examen”, astfel încât, practic, acest document să facă parte integrantă, activă din procesul educațional o perioadă de timp semnificativă pentru formarea individuală, evitându-se astfel elementele de stres și de „surpriză” care sunt asociate încă la noi, din păcate, examenelor.

Structura unei programe de examen elaborată de instituția specializată, responsabilă cu proiectarea și administrarea examenelor externe este, în general, următoarea:

- O introducere privind relația cu restul componentelor curriculare (reflecția criteriilor de evaluare conținute în curriculum).
- Scopurile generale ale programei de examen, specificând cerințele examenului, particularizate pentru o anumită disciplină de examen, precum și relația dintre componentele acestuia (de exemplu, examenul scris, oral, practic, componenta de evaluare a activităților curente, acolo unde este cazul).
- Exemple de obiective de evaluare și relația acestora cu obiectivele cadru și obiectivele de referință din curriculum.
- Matricea de specificații a examenului, ilustrând modul în care diferite obiective de evaluare vor fi reprezentate/evaluate în diferitele componente ale examenului.
- Schema de examinare, acolo unde examenul este realizat pe niveluri ale performanței așteptate, candidații având posibilitatea să opteze de la început pentru nivelul de bază sau cel extins.
- Conținuturile specifice disciplinei prezentate în manualele și auxiliarele valabile. Abilitățile, elementele de competență, atitudinile așteptate din partea candidaților pentru atingerea anumitor niveluri de performanță.
- Descriptori și/sau criterii de performanță care vor fi utilizabili la corectarea/aprecierea lucrărilor scrise sau în aprecierea răspunsurilor orale.
- Elemente/recomandări privind moderarea procesului de corectare/notare, precum și de apreciere orală, cu exemplificări care să funcționeze în procesul de asigurare a comparabilității judecăților de evaluare la nivelul tuturor evaluatorilor implicați în examenul respectiv.
- Calendarul examenului și elemente de regulament al desfășurării sale.

Modul în care este proiectată și utilizată programa de **examen** reflectă în bună măsură gradul de complexitate al examenului, de la stabilirea obiectivelor și până la structura testelor. Dintre opțiunile de viitor privind posibilitățile de diversificare a proiectării examenelor naționale menționăm introducerea nivelurilor. Reprezentând o tendință din ce în ce mai evidentă în cazul examenelor din Europa, nivelurile examenelor constituie introducerea mecanismului concret de a oferi tuturor elevilor șansa de a demonstra ceea ce pot face, în condițiile în care sunt asigurate în același timp și mijloacele de discriminare și de selecție necesare unui examen respectând standardele calitative.

Există mai multe modele posibile pentru proiectarea nivelurilor de dificultate a examenelor:

1. Modelul ofertei curriculare unice pentru toți candidații care se prezintă la examen și susțin o probă unică. Acestea i se poate adăuga o probă opțională, extinsă, pentru cei care doresc un bonus de punctaj, în baza demonstrării unor abilități extinse.

2. Modelul opțiunii pentru una dintre cele două probe posibile: una de nivel de bază și cealaltă de nivel avansat,

3. Modelul mixt, în care proba pentru nivelul de bază conține anumite întrebări/probleme de nivel mai ridicat iar proba pentru nivel avansat conține și anumite întrebări/probleme de nivel de bază, cu scopul de a asigura tuturor elevilor posibilitatea de a demonstra ceea ce pot face cu adevărat.

Introducerea **examenelor pe niveluri de dificultate și complexitate** implică nu numai o opțiune fermă de politică educațională, ci și asumarea responsabilității de asigurare a unui nivel profesionist obligatoriu în proiectarea examenului. În acest scop, trebuie activat **un mecanism de monitorizare a nivelurilor de dificultate de-a lungul timpului, astfel încât rezultatele** candidaților să poată fi privite ca fidele, iar aceștia să fie convinși că se vor pregăti și vor putea opta pentru nivelul de dificultate care li se adecvează cel mai bine. Introducerea nivelurilor de dificultate trebuie să fie optim corelată cu modul de certificare a abilităților, elementelor de competență și eventual, atitudinilor demonstrate de către candidat. În final, certificatul sau diploma primite trebuie să reflecte cu exactitate nivelul de dificultate, conferind „valoarea” de recunoaștere a întregului examen.

1.4. Proiectarea demersului de evaluare și examinare

În mod tradițional, la noi conceptul de **strategie** a cunoscut o carieră educațională extinsă, în strânsă relație cu teoriile instruirii, precum și cu aplicarea lor practică și teoretică — sintagma „**strategie de instruire**” desemnează: „un mod de combinare și organizare cronologică a ansamblului de metode și

mijloace alese pentru a atinge anumite obiective” (definiție UNESCO, din 1976); sau: „o ipoteză de lucru, o linie directoare de acțiune căreia i se asociază un anumit mod global de organizare a învățării și a condițiilor învățării, de utilizare cu precădere a unor metode și mijloace (I. Cerghit, 1988)”; o altă definiție, din perspectiva teoriei operaționalizate a acțiunii este: „o acțiune decompozabilă într-o suită de decizii — operații” fiecare decizie asigurând trecerea la secvența următoare de instruire, pe baza valorificării informațiilor dobândite în etapele anterioare.” (Dan Potolea, 1983).

Astfel, din perspectiva globală a procesului de formare/învățare/ instruire, strategia de instruire este un concept operațional extrem de util celor ce proiectează și gestionează în mod direct procesul. Caracteristica sa principală, precum și punctul său forte este calitatea de a armoniza în bază metodologică toate elementele componente. Într-o analiză critică a conceptului, Ioan Neacșu (op. cit., p. 219) sintetizează: „Astfel, putem socoti ca important faptul că strategia de instruire implică elevul în situații specifice de învățare, raționalizează și adecvează conținutul instruirii la particularitățile personalității elevilor (motivație, model de pregătire, stil de cunoaștere/de învățare), creează premise pentru manifestarea optimă a interacțiunilor dintre celelalte componente ale procesului de instruire, dependente, la rândul lor, de personalitatea profesorului, în special de capacitatea lui de a realiza proiectarea, implementarea și evaluarea instruirii”.

Conceptul de **strategie de evaluare desemnează o realitate care, din perspectiva procesului care se desfășoară în clasă, este o parte componentă integrantă, vitală prin funcția sa determinant reglatoare, a procesului de instruire. Prin dimensiunile sale „tehnice”, specializate, orientate spre un scop de evaluare clar stabilit și făcut transparent cu acuratețe, strategia de evaluare adoptată potențează procesul educațional în direcția dorită de cel care o proiectează și o aplică.**

În mod curent, în educația românească sunt semnalate ca funcționale două strategii de evaluare principale și larg utilizate:

a) Strategia de evaluare sumativă (sau cumulativă) care, datorită perioadei relativ îndelungate necesitate de prelucrarea eficientă a *feedback-ului* (i.e., rezultatele, estimarea performanțelor sub toate formele și de orice natură, ale elevilor), cere un timp de prelucrare a datelor mai îndelungat, imprimând astfel un ritm mai lent succesiunii secvențelor de evaluare.

Alte caracteristici importante ale strategiei sunt:

- caracterul „retrospectiv” al acțiunii evaluative intenționate;
- valoarea de diagnosticare pe termen mediu a informațiilor finale obținute în urma prelucrării datelor;
- caracterul globalizator, de însumare a unor elemente ale judecății de valoare, pe care îl are caracterizarea privind „succesul” sau „eșecul” în funcție de proba/probele de evaluare aplicate;

b) Strategia de evaluare formativă (sau continuă) care, prin oportunitatea de reglare a sistemului în mod frecvent și în pași mici, succesivi, asigură o periodicitate eficientă procesului în ansamblu.

Alte caracteristici ale strategiei sunt:

- este „implicită” în procesul educațional;
- are potențialul de a identifica atât punctele tari, cât și punctele slabe ale procesului;
- mizează pe valoarea „diagnostică” a judecății de valoare cu care se finalizează, stimulând și o posibilă analiză a mecanismelor și cauzelor eșecului sau succesului.

Atât pentru strategia **evaluării formative**, cât și pentru cea a evaluării **sumative**, punctul cheie îl reprezintă decizia asupra momentului în care are loc evaluarea reală: aceasta trebuie să permită evindetșierea, prin tehnicile și mijloacele de evaluare pentru care s-a optat, o schimbare relevantă în ceea ce privește comportamentul, rezultatele, atitudinile etc. ale celor evaluați.

În planul proiectării curriculare integrate, dihotomia **formativ/sumativ** (în forma sa cea mai cunoscută, cea introdusă de către Michael Scriven, în 1967, în articolul său, *The Methodology of Evaluation*, publicat în *Perspectives of Curriculum Evaluation*, nr. 1) trebuie analizată și aplicată în acțiunea educațională prin respectarea și adecvarea următoarelor principii:

1. Principalul scop al oricărei **acțiuni** de evaluare formativă este cel de a optimiza, prin analiza *feedback-ului* produs, anumite etape ale procesului educațional;

2. Principalul scop al oricărei **acțiuni de evaluare sumativă este cel de a evidenția efectele, eficiența, rezultatele globale ale procesului educațional, aflat într-un moment de final sau necesitând o decizie de schimbare;**

3. Cei doi factori ai dihotomiei pot funcționa cu maximă eficiență doar **în tandem**, potențându-se reciproc.

Așadar, atât la nivelul gândirii evaluării în planul proiectării curriculare, cât și din perspectiva integrării tuturor rezultatelor evaluării curente, realizată de către profesor, la clasă, logica evaluării necesită în primul rând asigurarea unei armonizări de substanță a demersului de evaluare cu cel general, educațional.

La nivelul procesului de examinare, cele două principale funcții ale examenelor — **funcția de certificare și funcția de selecție** — pot fi optim activate prin **două** modalități:

1. Clarificarea scopului principal al examenului, precum și a obiectivelor acestuia, în termeni transparenți pentru un număr cât mai larg de beneficiari;

2. Stabilirea și menținerea transparenței a unui raport între **certificare și selecție** ușor de explicat, de gestionat și de tradus în termenii obiectivelor probelor și instrumentelor de evaluare proiectate și aplicate.

Literatura de specialitate menționează puternice efecte de *backwash* în cazul actualizării ambelor funcții (pentru aprofundare, v. capitolul următor al prezentei lucrări). Diminuarea sau, după caz, „pozitivarea” acestui efect al examinării trebuie să constituie o țintă permanentă pentru cei ce proiectează, analizează și administrează examenele, precum și relațiile acestora cu sistemul educațional în general și cu domeniul extins al curriculum-ului în mod special.

Din perspectiva evaluării și a proiectării instrumentelor de evaluare conform unor standarde profesionale „acceptabile” în termeni de indicatori precum validitatea și **fidelitatea** — calități esențiale ale instrumentului de evaluare ca atare, ca și a performanței așteptate — problema practică principală pe care o ridică în acest moment la noi evaluarea curentă prin aplicarea/utilizarea unui instrument de evaluare este relaționarea sistematică și transparentă a celor trei componente:

a) abilitatea/abilitățile ce trebuie sau se doresc a fi evaluate prin instrumentul de evaluare;

b) elementul / elementele de conținut, ținând de specificul curricular, (de exemplu, tema concretă a unui eseu, în cazul disciplinelor umaniste);

c) criteriile de corectare/notare, care se materializează în grila sau baremul de corectare și notare și care vor fi, în final, transpuse în nota sau calificativul pe care le va primi elevul.

Aceste trei „componente critice” ale metodologiei evaluării, prezente, sub diferite forme și cu diferite accente și în practica din clasă, curentă a evaluării și examinării românești, au nevoie de o „transparentizare” în termenii explicitării im-pactului și efectelor conceperii, proiectării și aplicării lor în diferite situații de evaluare și, eventual, de examinare reale.

Vom analiza, în continuare, proiectarea **descriptorilor de performanță în** contextul românesc, precum și avantajele și dezavantajele oferite de introducerea **criteriilor calitative de notare**, concepte operaționale active și în sistemul nostru educațional în ultimii doi ani.

Sub denumirea de descriptori de performanță (sau descriptori de bandă) (engl. *band descriptors*) intră o categorie foarte diversă de formulări, descrieri, explicitări organizate în manieră ierarhică, ale performanței așteptate din partea elevilor aflați în situație de evaluare. Performanța așteptată este concepută ca performanța tipică, medie, realizabilă în primul rând în condițiile evaluării curente, de la clasă.

În sistemul de evaluare românesc, descriptorii de performanță au fost introduși la începutul anului școlar 1998-1999, la nivelul învățământului primar, ca un „act” al reformei educaționale în domeniu. În acest context, descriptorii stau la baza creării și implementării unui sistem complet nou de evaluare, respectând criterii unitare, aplicabile la nivel național, în aprecierea rezultatelor elevilor.

Până în acest moment, *feedback-ul* din partea învățătorilor, al inspectorilor, al elevilor și părinților, poate fi sumarizat astfel:

- evaluarea curentă a devenit mai unitară și, probabil mai obiectivă sub raportul consistenței procesului;

- faptul că sistemul descriptorilor de performanță nu este conceput să discrimineze puternic în sens normativ, între elevii aceleiași clase (spre deosebire de sistemul tradițional până acum la noi, prin notele școlare), s-a dovedit deja a fi benefic pentru acest nivel de vârstă;
- a apărut evidentă necesitatea complementării celor patru trepte ale notării elaborate în prima etapă, cu descrieri ale performanței de tip „excelent”, precum și a celei de tip „insuficient”;
- finalizarea judecății evaluative a învățătorului privind performanța elevului printr-un calificativ ce corespunde descrierii performanței demonstrate are implicații de profunzime, directe și asupra modalităților de înregistrare a rezultatelor școlare. Documentele care „certifică” aceste rezultate au intrat deja în sistem începând cu anii școlari ce au urmat introducerii calificativelor;
- introducerea concomitentă în sistem a noului curriculum pentru învățământul primar și a noului sistem de evaluare a rezultatelor școlare a avut avantajul evident de a potența importanța și impactul schimbării;
- principalele calități ale noului sistem de evaluare introdus sunt **flexibilitatea și adaptabilitatea sa: odată** introduse modelele-matrice pentru punerea în ecuația evaluativă a principalelor obiective de referință de la nivelul curriculum-ului, precum și a principalelor capacități, sistemul este capabil de a se autoregla.

Calitatea de „instrument de evaluare” a descriptorilor este ilustrată prin relația de directă proporționalitate între nivelurile de performanță identificate și descrierea calitativă a acestora, într-o grilă orientată crescător, de la nivelul inferior către nivelul superior. Modulurile în care acești descriptori de bandă pot fi creați, adaptați, utilizați în mod concret în activitatea de evaluare curentă și/sau de examinare este o chestiune care ține de activitatea evaluatorului/examinator, de efectele scontate și de importanța judecății de evaluare cu care se finalizează „aplicarea” descriptorului.

În acest context, devine esențială **armonizarea** metodologică și **pragmatică a parcursului obiective/instrument de evaluare/schemă de corectare/notare**. Structurate în termeni de performanțe și capacități urmărite, **obiectivele de evaluare** pot pregăti terenul pentru o schimbare curriculară ilustrând și dorita schimbare de mentalitate ce ar permite în sfârșit regândirea disciplinei în raport cu alte discipline, dintr-o perspectivă cu adevărat **transcurriculară**.

În acest scop, **instrumentele de evaluare adecvate** ar trebui să satisfacă în primul rând optima proporție dintre cantitativ și calitativ în procesul de evaluare/apreciere. Spre exemplu, un anumit tip de item — eseul structurat — poate fi completat printr-o schemă de corectare/notare sub forma unui descriptor exprimat printr-o scală de apreciere calitativă, analitică a nivelurilor abilității vizate.

În cadrul scalei, rubricii „notelor” îi poate fi asociată rubrica descrierii calitative a performanței. La fel, în cazul examenului oral, aprecierea răspunsului se poate face pe baza unei grile de evaluare modelată sub forma unui descriptor de performanță (v. exemplul proiectat și realizat de către S.N.E.E. pentru examenul de Bacalaureat 2000 la Limba și literatura română).

Descriptorul de performanță concretizat în forma unei **scale de apreciere** are avantajul practic de a argumenta judecata de valori, întemeind-o în mod sistematic. Acest lucru pare a fi extrem de util în evaluarea curentă, internă, realizată de către profesor la clasă, chiar în condițiile în care necesarul de timp acordat întregului proces de apreciere este relativ important.

Spre deosebire de aplicarea/experimentarea unui **instrument de evaluare** de tip test (de cunoștințe, de performanță, de abilități etc.) sau chiar a unui portofoliu spre exemplu (ca instrument de evaluare complementară), în care numărul și rezultatele subiecților implicați sunt esențiale pentru întemeierea unor concluzii relevante și credibile, aplicarea **descriptorului de performanță** sub forma **scalei de apreciere** descriind niveluri de performanță așteptată corespunzătoare unor **note-prag** sau, eventual, unor calificative, este un demers ce vizează în primul rând reacții ținând de utilizarea directă, la clasă, de către practician.

Principalele argumente pentru introducerea, adaptarea și utilizarea unor asemenea instrumente calitative în cazul aprecierii și notării curente sunt:

- potențialul lor formativ în termeni de înțelegere și modelare. a așteptărilor în ceea ce privește performanța dorită;
- stimularea capacităților metacognitive și de autoevaluare a demersului individualizat, a performanței, a procesului de învățare în general, ca și al celui de evaluare în mod specific;

- creșterea transparenței și a consistenței procesului de apreciere/notare și prin aceasta, a credibilității deciziei, a judecății de valoare în cazul evaluării curente, cu impact direct în perspectivă și asupra examenelor;
- întemeierea, în timp, a consensului de tip „profesionist” între cei care acționează și ca evaluatori în activitățile curente, și ca examinatori, în situația de examen în mod special — profesorii și învățătorii. Realizarea acestui tip de consens este esențială pentru optimizarea continuă a calității judecății experte, bazate pe criterii asumate, transparente și operaționalizabile în mod concret, răspunzând nevoilor reale;
- stimularea, și pe această cale, deși într-un mod indirect, dar „puternic” — prin utilizare consecventă și „adaptată” — a schimbării accentului și ponderii de pe conținuturi curriculare pe competențe și capacități activate contextual sau situațional, unde conținuturile sunt cu adevărat ilustrative și nu determinate.

1.5. Constrângeri și riscuri în proiectarea demersului de evaluare și examinare

Demersul de evaluare și examinare prezintă o logică internă ce îi conferă specificitate și identitate proprie. Proiectarea demersului conform acestei logici garantează acțiuni de evaluare aparținând la domeniul științificului, conform unei etici profesionale care constituie cadrul comun a ceea ce poartă numele de paradigma **evaluării educaționale**. În ultima decadă, în culturile educaționale vestice, aflate într-o schimbare profundă de la ceea ce poartă numele de „o cultură a testării” (engl. a *testing culture*) spre ceea ce se dorește a fi „o cultură a evaluării” (engl. *an assessment culture*), domeniul evaluării și examinării educaționale este supus tendinței majore de limitare a impactului viziunii strict psihometrice prin ponderare cu utilizarea din ce în ce mai extinsă a instrumentelor de evaluare complementare sau alternative și a modurilor calitative de tratament și interpretare a datelor rezultate.

În acest context, sistemele educaționale în care **cultura evaluării educaționale** abia începe — prin reforma educațională structurală — să înglobeze coordonata psihometrică într-o măsură cu adevărat semnificativă, se află în situația de a beneficia de experiențele pozitive, de succesele înregistrate până acum, dar și de a se feri de exagerările drumului deja parcurs de către alte sisteme educaționale.

Constrângerile proiectării demersului de evaluare sunt date de natura sa duală: **proces științific** prin demersul necesitat, dar și **social** prin determinare și necesitatea adecvării — cu cerințe ținând de deontologie, dar și de costuri — ca activitate publică, dispunând de un buget public ce necesită responsabilizarea unor factori participanți, precum și o completă transparență în comunicare, datorată tuturor celor implicați sau vizați.

În proiectarea oricărui demers de evaluare sau de examinare există o serie de **riscuri pe care le enumerăm succint**:

- riscul formulării unor scopuri și obiective neclare sau neprecizate în termeni concreți de la bun început — falsifică întregul proces, dezorientând așteptările celor direct implicați;
- riscul nepotrivirii între scopuri și modurile evaluării (ratarea condiției de „adecvare la scop”) — conduce la neconcordanțe flagrante în fluxul, în succesiunea logică a etapelor evaluării sau examinării;
- riscul neadecvării tehnicilor de evaluare la modurile enunțate — conduce la diminuarea drastică a eficienței întregului proces;
- riscul neadecvării tehnicilor de evaluare alese la scopurile și obiectivele enunțate — conduce la imposibilitatea de a produce date cu adevărat relevante privind abilitățile sau elementele de competență ale celor investigați în situația de evaluare/examinare și, în consecință, la eronarea judecății evaluative;
- riscul lipsei totale de comunicare sau al comunicării neadecvate a rezultatelor/datelor/concluziilor privind procesul de evaluare — conduce la imposibilitatea realizării *feedback-ului* și la diminuarea considerabilă a impactului procesului de evaluare atât asupra beneficiarilor, cât și asupra celorlalți participanți sau implicați în proces.

În proiectarea oricărui proces de evaluare sau de examinare, autoritatea **responsabilă** se confruntă cu situația de a trebui să îmbine cei mai potriviți factori educaționali cu cele mai acceptabile condiții administrative și financiare posibile, în timpul stabilit a fi dedicat procesului (de cele mai multe ori, unul redus, deoarece evaluarea este încă privită, din păcate, ca un proces concurențial cu practica

educațională „obișnuită”, curentă!). În plus, de vreme ce evaluarea sau examinarea își au propriile scopuri și obiective dominante, acestea trebuie să se situeze în perfectă armonie cu practica educațională de bună calitate, stimulând-o și promovând-o continuu.

Sumar de idei

Conceptele-cheie:

- evaluarea educațională;
- reforma sistemului de evaluare;
- evaluarea de proces, evaluarea curentă continuă a activităților zilnice;
- evaluarea națională;
- relația evaluare curentă-examen;
- mizele examenelor naționale;
- programa de examen;
- dihotomiile formativ/sumativ;
- continuu/final;
- strategii de evaluare formativă și sumativă;
- funcțiile de certificare și de selecție; descriptorii de performanță.

Idei principale:

- Evaluarea educațională este un subsistem critic al sistemului educațional, aflat la noi într-o reformă structurală, comprehensivă;
- Parcursul obiective instituite — experiențe de învățare — evaluare trebuie să fie unul coerent și suficient sieși;
- Relația evaluare curentă — examene trebuie să fie ponderată la nivel național, în sensul creșterii calității judecății evaluative în baza evidențelor de natură empirică;
- Examenele naționale se vor constitui într-un demers cu adevărat profesionist în momentul optimizării transparenței sale;
- dihotomia formativ/sumativ poate fi operaționalizată la nivelul proiectării curriculare prin respectarea a trei principii de bază;
- strategiile de evaluare formativă și sumativă sunt cele mai frecvent aplicate în evaluarea curentă, de la clasă;
- funcția de certificare a unui examen asigură recunoașterea finalizării unui parcurs educațional, în baza unor elemente clar specificate;
- funcția de selecție a unui examen asigură ierarhizarea, în bază criterială explicită, a celor examinați, în funcție de performanțele evidențiate;
- descriptorii de performanță reprezintă explicitarea în termeni calitativi a unor niveluri-cheie ale performanțelor tipice, medii sau ridicate ale celor evaluați sau examinați;
- criteriile de notare reprezintă descrieri în termeni calitativi ale performanțelor așteptate, asociate cu anumite sarcini de lucru, pentru anumite note sau intervale de notă;
- scalele și baremele de corectare și notare constituie părți componente esențiale ale instrumentelor de evaluare și examinare, de a căror acuratețe tehnică depind, în mare măsură, validitatea și fidelitatea procesului de evaluare în ansamblu.

Capitolul II. Relațiile dintre curriculum și evaluare

Evaluarea educațională trebuie privită ca o parte integrantă, activă și importantă a curriculum-ului, dezvoltând relații specifice și semnificative în termeni de impact și efecte asupra elevilor, asupra profesorilor, asupra celorlalți agenți educaționali, ca și asupra factorilor de decizie responsabilizați în cadrul procesului.

În România, după 1990, reforma curriculară a cunoscut mai multe etape de desfășurare și sistematizare. Conceptul central, cel de **Curriculum Național**, a devenit activ și operațional începând cu 1997, când concepătorii de curriculum și-au concentrat eforturile de ridicare a gradului de coerență în implementarea deciziilor de politică educațională vizând schimbări pe termen scurt, mediu și lung.

2.1. Efectele educaționale

Indiferent de nivelul la care se poziționează, orice act de evaluare care se finalizează printr-o judecată de valoare sau de existență cu impact asupra elevului aflat la orice al parcursului său educațional, participă într-o formă sau alta la furnizarea de informații privind anumite secvențe ale procesului educațional, efectele asupra elevilor sau rezultatele concrete. Aceste informații pot avea o natură, o structurare, forme și utilizări extrem de variate. Pe baza analizei acestor informații pot fi definite două tipuri de efecte ale acțiunii relației **curriculum - evaluare**:

a. Efectul de feedback

Circulația **fluxului de informație** obținută pe baza aplicării metodelor și tehnicilor de evaluare specifice, dinspre subiecții evaluării/examinării, către toți cei implicați, la un nivel sau altul, sub o formă sau alta, în procesul de evaluare, este o formă de *feedback* semnificativ deoarece în baza acestei circulații se pot lua deciziile dorite de optimizare a procesului. Monitorizarea acestui efect conduce la reglarea/optimizarea/ajustarea programului de formare, a cursului sau a activităților de predare/învățare/instruire. Pentru o „exploatare” optimă a efectului de către profesor răspunsul, comentariul, reacția informată, decizia acestuia trebuie să se succedă și să se finalizeze cu rapiditate.

Cea mai cunoscută formă de *feedback* este cea a fluxului informațiilor specifice dinspre elevi înspre profesori, pe două căi: prin intermediul evaluării curente a activităților zilnice și prin intermediul examenelor. Această mișcare a informației se face în spirală, în cadrul circuitului educațional.

Feedback-ul formativ (pentru detalii vezi și N. Groulund, 1981, p. 489-490), obținut în urma evaluării continue, contribuie în măsură determinantă la urmărirea și evidențierea progresului educațional al elevilor. Prin diagnosticarea deficiențelor, a punctelor slabe, ca și a punctelor tari ale educabilului, în cadrul procesului de evaluare curentă, continuă, dialogul educațional dintre profesor și elev capătă consistența și forța necesare pentru a schimba atitudinile. Calitatea comunicării în relația profesor-elev este esențială, iar *feedback-ul* formativ are atât menirea de a ridica nivelul motivațional al educabilului, cât și de a induce și întări în timp, comportamente metacognitive cu miză formativă majoră.

Feedback-ul sumativ, realizat în urma unei evaluări cu caracter preponderent sumativ, este un prețios instrument pentru profesor, deoarece îi furnizează acestuia datele necesare pentru a-și modifica, transforma, adapta demersul educațional viitor în funcție de răspunsurile la sarcină ale educabililor. Pentru elevii confrunțați cu evaluarea sumativă, acest tip de *feedback* are în primul rând semnificația recunoașterii valorii sau efortului personal, a „recompensei” de natură simbolică pe care trebuie să o primească toți cei care se află în situația de evaluare sau de examinare.

Rolul pe care poate să-l joace un *feedback* „pozitiv” într-un context educațional activ poate fi ilustrat prin demonstrația lui Royce Sadler, din articolul său *Formative Assessment: revisiting the Territory*. Astfel, *feedback-ul* este privit ca un posibil model pentru practica educațională de la clasă, în condițiile în care este activată funcția profesorului ca evaluator. La întrebarea „Ce aduce un profesor competent în actul evaluării?” sunt enumerate șase categorii de elemente:

- cunoaștere (sau cunoștințe) amplă/e și extinsă/e privind conținutul sau substanța a ceea ce trebuie învățat. Acestea pot fi de natură factuală, procedurală sau de aprofundare a unui anumit domeniu sau a unei discipline de studiu;
- un set de aptitudini, atitudini și dispoziții față de procesul de învățare, inclusiv „atitudinea de a empatiza cu cel ce învață”, culminând cu adecvarea demersului de evaluare curentă, internă, la caracteristicile și cerințele individuale;

- abilitatea de a construi / a crea / a redacta (sau de a selecta/alege) teste de evaluare adecvate sarcinilor de evaluare alese sau proiectate;
- cunoașterea extinsă și în profunzime a criteriilor și standardelor de evaluare potrivite sarcinilor de evaluare alese;
- expertiză în formularea judecăților de evaluare, pe baza experiențelor anterioare de apreciere a performanței elevilor în relație cu anumite sarcini;
- expertiză în comunicarea *feedback-ului* către elevi, părinți și, eventual, autoritățile educaționale.

În mod evident, ultimele două categorii de elemente pot fi satisfăcute doar de către profesorii cu experiența practicii evaluative, dar nu și cu cea a rutinei din domeniu. Autorul citat consideră că structura clasică a *feedback-ului* (caracterizat prin: acuratețe, comprehensivitate și adecvare) trebuie să fie completată cu ceea ce numește „valoarea catalitică și de antrenare, precum și cu abilitatea de a inspira încredere și speranță” (op. cit., p. 81) celor evaluați. În analiza acestui efect, dezbaterile se centrează din ce în ce mai mult pe calitatea sa, în sensul maximei sale pozitive.

b. Efectul de backwash (sau de washback)

Definit succint prin impactul vizibil, de substanță și de durată pe care îl au evaluarea și examinarea asupra curriculum-ului ca întreg, la modul general, și asupra disciplinei de studiu în mod particular, asupra zonei dintre ariile curriculare, precum și chiar și asupra schemei de evaluare în mod particular, acest efect poate acționa în mod pozitiv sau în mod negativ în cadrul procesului educațional.

În primul rând, **efectul retroactiv** sau **de siaj** cum ar putea fi tradus la noi, acționează de fapt asupra **curriculum-ului „ascuns”**, realizat în practică și nu, în mod direct, asupra celui oficial. În mod concret, nivelurile la care se poate observa acțiunea acestuia sunt:

- metodele curente de predare-învățare activate în clasă;
- nivelurile motivaționale ale elevilor;
- calitatea interacțiunilor directe dintre profesor și elev, inclusiv stimularea, prin evaluare, a conformării automate, în cazurile anumitor elevi, la un anumit tip de cerințe ale profesorului;
- nivelul strategiei unității de învățământ, care poate sugera, în baza rezultatelor diferitelor tipuri de evaluare sau a examenelor, preluarea de către anumiți profesori/învățători, a anumitor clase de elevi, „selectate” etc.

Deși se înregistrează un anumit **efect retroactiv** și ca urmare a evaluării curente, interne, **examenelor naționale**, precum și **evaluările naționale** sunt recunoscute a avea cele mai puternice efecte de această natură în cadrul procesului educațional.

Ponderarea în sens pozitiv a **efectului de backwash** se poate realiza pe mai multe căi și în mai multe moduri:

- prin „apropierea” cât mai mare a curriculum-ului evaluat de curriculum-ul oficial, în condițiile urmăririi unei acoperiri curriculare cât mai extinse în/prin situațiile de evaluare/examinare;
- prin introducerea continuă de noi metode, tehnici și instrumente de evaluare, în special a celor cu caracter alternativ sau complementar, furnizând date de natură calitativă și criterială privind performanțele vizate ale educabililor;
- prin asocierea și combinarea de instrumente de evaluare a căror aplicare să conducă la obținerea de date de natură cât mai diversă, interpretabile conform unei scheme de evaluare complexe, dar transparente, îmbinând normativul și criterialul deopotrivă;
- prin utilizarea continuă a efectului de *feedback* în sensul fluentizării procesului de evaluare/examinare;
- prin asigurarea nivelului de transparență maxim necesar implicării în proces — prin comunicare eficientă și adecvată — a tuturor celor interesați: elevi, părinți, factori de decizie responsabili, opinie publică;
- prin alocarea / identificarea de către autoritățile responsabile, a resurselor umane, materiale și financiare necesare desfășurării unui demers de evaluare sau de examinare eficient.

Efectul de siaj, cum mai poate fi numit (prin asociere metaforică a efectului evaluării sau examinării cu un vas navigând pe o mare liniștită, lăsând în urma sa cele două linii de valuri care pot reverbera

la distanțe apreciabile), are și o dimensiune predictivă, motiv pentru care literatura de specialitate înregistrează și conceptul de **washbefore**. Referindu-se în esență la efectele pe care evaluarea și examinarea le au și asupra curriculum-ului ulterior sau asupra parcursului educațional din etapa următoare, acesta are, din păcate, un impact mai puțin direct și vizibil asupra proiectării demersurilor de evaluare și examinare.

2.2. Standardele educaționale naționale și rolul lor în stimularea evaluării și examinării

Conceptul de **standard educațional** a devenit din ce în ce mai des utilizat, cu o varietate de sensuri și într-o multitudine de contexte mai ales odată cu ridicarea nivelului de profesionalism al dezbaterilor privind reforma educațională. Dintre categoriile de standarde educaționale cele mai cunoscute, menționăm: standardele curriculare de conținut, de predare, standarde de evaluare, standarde de examinare, standarde de formare profesională. În esență, fiecare aspect sau latură a procesului educațional poate fi acoperit prin formularea de standarde dezirabile, la care se raportează progresul sau pe baza cărora se poate stabili *feedback-ul*.

Standardele educaționale trebuie clar separate, datorită rolului lor, precum și datorită ariei lor de aplicare, de **standardele ocupaționale**, care nu fac obiectul dezbaterilor din domeniul educației.

Din perspectiva evaluării și examinării, **standardele de performanță** vizează performanțele elevilor în situația de evaluare sau de examinare, se află la interfața dintre curriculum și evaluare/examinare și, în esență, sunt formulări, în termen de comportamente observabile, a ceea ce elevii știu și pot să facă în contextul curriculum-ului parcurs. Acestea descriu ceea ce trebuie să fie atins — realizat — performat de către elevi după parcurgerea unei etape de formare și acționează ca „ținte” cunoscute atât de către elevi, cât și de către profesori, de urmărit în demersul lor comun. Spre exemplu, în Curriculum-ul Național britanic, elementele care fac fuziunea dintre evaluare și celelalte componente curriculare sunt așa-numitele *attainment targets*, (tr. aprox., sarcina-țintă de realizat), descriind atât cunoștințele, cât mai ales abilitățile capacitățile și competențele necesare a fi demonstrate de către elevi pentru a **atinge** sau a **realiza standardele de performanță** necesare, prescrise prin Curriculum Național. În ceea ce privește nivelurile la care sunt formulate, acestea pot defini **nivelurile de performanță minim acceptabile**, pe cele **ale performanței tipice** (sau **medii**), și pe cele **ale performanței optime sau de excelență**.

Deși sunt formulate în termeni calitativi, **standardele de performanță** trebuie să se bazeze, într-un anumit stadiu al stabilirii și formulării lor, pe **date cantitative reale, valide și fidele**. Aceste date pot proveni din două surse de bază: evaluările naționale și examenele naționale. Cheia formulării profesionale a acestor standarde stă în modalitatea de interpretare a datelor de bază, în funcție de principiile asumate inițial.

În continuare, oferim o comparație, din perspectiva evaluării și examinării, a modului în care sunt formulate standardele de performanță în două sisteme educaționale cu tradiție în acest domeniu — Marea Britanie și Olanda.

Marea Britanie	Olanda
<p>Curriculum-ul Național cuprinde:</p> <ul style="list-style-type: none"> • disciplinele prescrise, de bază (engl. <i>core</i>) și cele „din extensie” (engl. <i>foundation</i>) • programele de studii • sarcinile - țintă de atins în cazul fiecărui "domeniu" pentru fiecare disciplină • standardele sau formulările „nivelurilor” de realizare sau de performanță • nivelurile - țintă pentru elevii cu performanță tipică la finalul stadiilor - cheie(vârstele de 7,9,11,14 ani). • Performanțele elevilor sunt raportate la Curriculum-ul Național prin intermediul Evaluărilor Naționale la nivelul fiecărui stadiu - cheie. • Rezultatele Evaluărilor Naționale sunt comunicate în mai multe moduri : • rezultatele individuale sunt raportate elevilor și părinților • rezultatele globale pe școli sunt făcute publice • tabelele normative conținând rezultatele școlii(engl. <i>league tables</i>) sunt publicate. 	<p>Cadrul Educației de bază cuprinde:</p> <ul style="list-style-type: none"> • obiective transcurriculare • sarcina - țintă specifică pe disciplină • monitorizarea și evaluarea, precum și optimizarea sistemului <p>Evaluarea Națională a Progresului educațional oferă:</p> <ul style="list-style-type: none"> • descrieri a ceea ce elevii știu și pot să facă • informații detaliate privind stăpânirea /realizarea sarcinilor-țintă • date empirice pentru evaluarea și optimizarea sistemului. <p>Standardele "relative" se stabilesc la nivel "minimum": ceea ce poate fi realizat de aproximativ 80% dintre elevi; și la nivel ridicat: ceea ce poate fi stăpânit/realizat de către cei mai buni 40% dintre elevi.</p> <p>Standardele sunt stabilite în urma unui proces în trei etape:</p> <ul style="list-style-type: none"> • Stabilirea standardului(pe baza judecății și a unui număr de itemi care vor fi organizați în teste și aplicații) • discuții de grup(în paneluri care stabilesc standardele preliminare la nivel minim și la nivel ridicat) • analiza discrepanțelor(revizuirea standardelor preliminare se face pe baza datelor oferite de performanța reală).

Tabelul nr. 1 - Comparație între modurile de formulare a standardelor de performanță în Marea Britanie și Olanda

La noi, apariția „**Curriculum-ului Național pentru învățământul Obligatoriu. Cadru de referință**” semnaleză apariția proiectului de realizare a standardelor de ciclu curricular ca „nivel intermediar de standardizare”. În terminologia adoptată, programele școlare sunt elaborate pe discipline și cuprind:

- „modelul curricular al disciplinei de la clasa I la clasa a VIII-a (engl. „*curriculum chart*”)
- obiectivele cadru ale disciplinei pentru învățământul obligatoriu (engl. „*content/skills domains*”)
- obiectivele de referință (engl. „*attainment targets*”)
- activitățile de învățare recomandate (engl. „*recommended learning activities*”)
- conținuturile sugerate pentru autorii de manual (engl. „*content*”)
- recomandări adresate profesorului, privind acțiunea didactică și elaborarea curriculum-ului la decizia școlii, a temelor cross-curriculare (engl. „*recommended teaching activities/suggestions concerning school-based curriculum development*”)
- standardele de performanță pentru finele celor trei cicluri curriculare discutate anterior: ciclul achizițiilor fundamentale, ciclul de dezvoltare, ciclul de observare, orientare (engl. „*performance standards*”)(op. cit., p. 39).

Am citat fragmentul de mai sus pentru a facilita evidențierea următoarelor observații:

- diferența este vizibilă între cele două tipuri de programe: programa școlară (sau de studiu, sau analitică) și programa de examen;
- sarcinile-țintă de atins ce funcționează în curriculum-ul britanic spre exemplu, sunt formulate în curriculum-ul românesc în termenii obiectivelor de referință;
- standardele de performanță curriculare (deci definite din perspectiva curriculum-ului) sunt privite ca standarde de final de ciclu curricular, sunt formulări ale unor niveluri medii de performanță, au în vedere finalitățile pe ciclu și treaptă de școlaritate, obiectivele cadru și pe cele de referință ale obiectului de studiu, precum și elemente de psihopedagogia vârstelor adecvate fiecăruia;
- în momentul proiectării lor, aceste standarde curriculare au în vedere **curriculum-ul oficial** (cel stabilit prin intermediul documentelor cu caracter reglator). În momentul revizuirii lor, după o perioadă de „implementare a curriculum-ului”, standardele vor putea include și elemente aparținând **curriculum-ului implementat**, „ascuns”, care implică și reflectă învățarea/formarea de reprezentări mentale, atitudini, norme, concepții, convingeri „formalizate” ca ritualuri sociale, reguli și norme identificate în realitatea educațională.

Din perspectiva evaluării și examinării, **standardele de performanță** sunt formulări care reflectă următoarele principii:

- au în vedere **performanța realizată**, demonstrată, „activată” în și prin situația reală de evaluare și de examinare;
- reflectă acea parte a curriculum-ului care a fost atinsă „cu adevărat” și într-un mod relevant, capabil „să lase urme” în mentalul elevului/educabilului, pe parcursul procesului său de formare;
- reflectă ceea ce elevul știe și poată să facă sau este capabil să comunice despre ceea ce poate să facă;
- echilibrul dintre nivelul de generalitate și cel de specificitate al formulării acestor **standarde de performanță** este dat de relația specifică dintre **standarde** și **obiectivele de evaluare**, fiecare sistem educațional impunând propria sa viziune în această privință.

În ceea ce privește **standardele examenelor**, în special cele ale examenelor naționale externe, proiectate și administrate de o agenție sau de către un serviciu specializat, una dintre cele mai importante mize pentru acestea este **recunoașterea valorii** sau **a creditelor** reflectate la nivel internațional. Crearea pe bază profesionistă a examenelor naționale implică asigurarea și monitorizarea calității examenelor pe tot parcursul proiectului, administrării și raportării rezultatelor acestora, prin crearea unor mecanisme de control al calității recunoscute și acceptate și pe plan inter-național. Unu[dintre aceste mecanisme este aderarea instituției organizatoare la standardele internaționale, în cazul de față, **Standardul ISO 9001** un standard de apreciere a calității examenelor. Această aderare înseamnă de fapt asumarea responsabilității majore de a urma procedurile implementării standardului pe o perioadă „de probă”, precum și de asigurare a unei comunicări „transparente” a etapelor, rezultatelor, efectelor etc. ale întregului proces.

2.3. Obiective cadru, obiective de referință, obiective de evaluare. Operaționalizarea obiectivelor din perspectiva evaluării și examinării. Rolul taxonomiilor

Curriculum-ul național românesc operează cu câteva categorii de obiective: **obiectivele ciclurilor curriculare, obiectivele cadru și obiectivele de referință**. Dacă prima categorie funcționează, din perspectiva conectorilor de curriculum, după cum am văzut, ca standarde curriculare de performanță, celelalte două categorii reflectă două niveluri de generalitate diferite în cadrul aceleiași discipline de studiu, pe parcursul mai multor ani de studiu cele „cadru”, și de la un an la altul, cele „de referință”. Toate vizează „formarea unor capacități și atitudini specifice disciplinei”, precum și „progresia în achiziția de competențe și de cunoștințe de la un an de studiu la altul” (*Curriculum Național*, op. cit., p. 39).

În acest context, obiectivele **de evaluare** sunt obiectivele având acel grad de specificitate care permite o **măsurare educațională** caracterizată printr-un grad suficient de obiectivitate pentru a fi validă și fidelă, deoarece acest lucru se face pe baza aprecierii, cu ajutorul unui instrument de evaluare, a comportamentului observabil al celui evaluat, aflat în situația de evaluare sau de examinare. La rândul lor, instrumentele de evaluare sunt create pe baza unor obiective cu un grad de

specificare mai ridicat (sunt obiective operaționalizate, răspunzând celor trei întrebări clasice: **ce se evaluează?**; **cum se evaluează?** și **cât se evaluează?**), facilitând „măsurarea” performanțelor.

În continuare, prezentăm o diagramă a procesului de proiectare a obiectivelor de evaluare (preluată după Macintosh, op. cit., p. 14). Accentul cade pe relația care conduce la satisfacerea scopurilor propuse, iar exemplul este interesant datorită perspectivei globalizatoare asupra întregului proces.

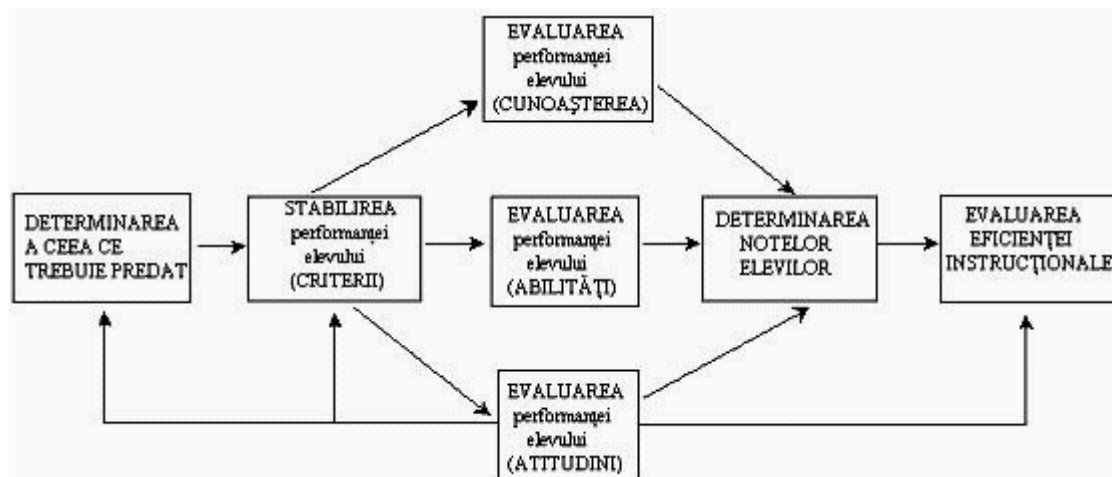


Figura nr. 1 - Diagrama procesului de proiectare a obiectivelor de evaluare (după Macintosh)

În articolul său de referință, *A Twenty-Year Perspective on Educational Objectives*, W.J. Popham face o analiză de profunzime a dezbaterii privind obiectivele comportamentale în educație, din două perspective — cea strict istorică, a „vocilor” care au adus ceva semnificativ în dezbaterile privind obiectivele; și cea analitică, a deconstrucției „mitului” formulării obiectivelor educaționale.

O primă concluzie a analizei perioadei 1960-1970 a fost următoarea: datorită tendinței de a formula cât mai multe obiective care să acopere cât mai complet plaja comportamentelor elevilor, s-a ajuns la „atomizarea” acestora într-un mod care le face inutile din perspectiva unei evaluări comprehensive. Soluția sugerată pentru a ieși din impasul pierderii din vedere a pădurii de dragul copacilor: „În loc de a se concentra asupra a 40 de obiective cu scop restrâns, evaluatorii le vor prezenta factorilor de decizie doar 5-6 obiective măsurabile mai larg formulate” (op. cit., p. 190).

O altă concluzie vizează gradul de utilitate reală a taxonomiilor obiectivelor comportamentale. Popham atrage atenția asupra faptului că taxonomiile „vizează procese individuale ascunse, acoperite, a căror natură trebuie să fie dedusă din comportamente vizibile, observabile” (op. cit. p. 191). Așadar, gradul de incertitudine este ridicat prin interpunerea multiplelor instanțe interpretative, ceea ce impune o abordare precaută a relației obiectiv - grad de performanță estimat.

De asemenea, de importanță majoră este observația privind natura reală a acestei relații: abia 20 de ani mai târziu au aflat cercetătorii că, de fapt, comportamentul identificat în obiectiv și nivelul de performanță căutat în /prin acel comportament sunt separabile în mod decisiv” (op. cit., p. 193). Este recomandabil, prin urmare, ca acestea să fie formulate separat — dar în termeni comparabili! — deoarece „amestecarea aspirațiilor privind comportamentele cu standardele de performanță în cadrul obiectivelor educaționale nu aduce claritate, ci confuzie” (id. ibid.).

Concluzia finală a articolului citat reflectă poziția teoriei și practicii actuale în domeniu: „În evaluările educaționale curente, rolul obiectivelor educaționale este tipic modest. În vreme ce mai înainte evaluatorii educaționali erau în mod frecvent prinși de importanța formulării obiectivelor, evaluatorii educaționali de astăzi se concentrează în mod obișnuit asupra datelor privind efectele programului, indiferent dacă aceste efecte au fost „asigurate” prin obiectivele prespecificate” (id. ibid.). În consecință, Popham oferă cinci recomandări pentru evaluatorii care operează cu obiective educaționale:

- „Evaluatorii educaționali trebuie să formuleze sau să recomande obiective educaționale astfel încât gradul în care un obiectiv este atins să poată fi determinat în mod obiectiv.

- (...) trebuie să evite formularea de numeroase obiective înguste și, în schimb, să se concentreze asupra unui număr rezonabil de obiective cu scop mai larg, deoarece evaluarea este un efort orientat către decizie.

- (...) trebuie să utilizeze taxonomiile doar ca pe niște instrumente euristice orientative, nu ca pe niște instrumente analitice rafinate.

- Dacă instrumentelor de măsură li se cere să ateste atingerea unor obiective educaționale, evaluatorii educaționali trebuie să utilizeze metode și instrumente pe cât posibil criteriale și nu normative.

- Evaluatorii educaționali trebuie să vizeze în mod separat accentul pe comportamentul enunțat în obiectivul educațional și nivelurile de performanță așteptate din partea elevilor." (op. cit., pp. 193-194).

Așadar, perspectiva evaluării pune accente diferite în procesul de formulare a obiectivelor față de demersul proiectării procesului instructiv-educativ. Punctul de plecare îl reprezintă, desigur, tot taxonomiile obiectivelor comportamentale, bine cunoscute și în spațiul românesc:

- Taxonomia lui Bloom et. al., apărută în 1956 (*Taxonomy of Educational Objectives: Handbook 1: the Cognitive Domain*);

- Taxonomia lui Krathwohl et. al., apărută în 1964, vizând, în completarea cele bloomiene, cele cinci niveluri al domeniului afectiv;

- Taxonomia lui Guilford, Alfred de Block, modelele lui R. Gagne și Gerlach & Sullivan etc.

Din perspectiva evaluării, taxonomia reprezintă cadrul ce stă la baza construirii obiectivelor de evaluare. Categoriile conținute pot fi cele clasice, bloomiene:

- Cunoaștere

- Comprehensiune sau înțelegere

- Aplicare

- Analiză

- Sinteză

- Evaluare

Un exemplu mai puțin cunoscut la noi, dar important deoarece se bazează pe categorii de întrebări folosite în testare în mod curent în Statele Unite, este **Taxonomia propusă de Ebel** în 1979 și citată de Macintosh (op. cit., p. 15):

1. „Înțelegerea terminologiei (sau a vocabularului).

2. Înțelegerea faptelor și principiilor (sau generalizarea).

3. Abilitatea de a explica sau de a ilustra (înțelegerea relațiilor).

4. Abilitatea de a calcula (probleme de factură numerică).

5. Abilitatea de a face predicții (ceea ce este probabil a se întâmpla în condiții specificate).

6. Abilitatea de a recomanda acțiunea potrivită (în cazul anumitor probleme practice specifice).

7. Abilitatea de a face o judecată evaluativă."

Pe baza categoriilor de acest fel, sunt generate obiectivele de evaluare, prin definirea comportamentului observabil, folosind lista binecunoscutelor verbe adecvate. Procedura de operaționalizare cea mai cunoscută și mai frecvent folosită, datorită simplității sale, este cea propusă în 1962 de către Robert Mager, în manualul său, *How to write Instructional Objectives*.

Astfel, formularea obiectivului de evaluare trebuie să satisfacă **toate** stagiile sau etapele următoare:

- să identifice, numind comportamentul vizat, pe care elevul trebuie să-l demonstreze;
- să identifice cu claritate condițiile importante în care comportamentul se poate produce sau poate deveni vizibil, măsurabil;
- să precizeze un nivel al performanței acceptabile, prin enunțarea unui criteriu de reușită direct măsurabil.

Din perspectiva evaluării și examinării educaționale, există o diferență de esență între formularea (și, eventual, operaționalizarea obiectivelor procesului de predare-învățare, realizată de cele mai multe ori de către profesor, în cadrul planificărilor anuale, semestriale și zilnice) și operaționalizarea obiectivelor ca etapă esențială în crearea unui instrument de evaluare — acțiune specifică activității de evaluare sau de examinare.

Obiectivele care stau la baza creării unui instrument de evaluare, indiferent de natura acestuia (standardizat sau creat de profesor; folosit la evaluarea formativă sau sumativă etc.), trebuie să fie clar formulate în termeni operaționali, astfel încât să fie evident faptul că sunt măsurabile cu eficiența dorită. Obiectivele sunt un element-cheie în construirea instrumentului de evaluare, astfel încât calitatea și acuratețea operaționalizării lor sunt esențiale.

În continuare, oferim un exemplu de operaționalizare la disciplina Limba și literatura română. Primele două obiective (cel „cadru” și cel „de referință”) sunt preluate din documentul oficial în uz la clasa a IX-a și au fost alese pentru a ilustra succesiunea normală în care trebuie să se situeze cele trei categorii de obiective. Obiectivul de evaluare formulat precizează **ce, cum și cât** trebuie să fie capabil să realizeze sau să performeze elevul. El poate funcționa, ca atare, atât la finalul clasei a IX-a, cât și ca obiectiv al probei scrise pentru examenul de Bacalaureat.

Obiectivul cadru:

1. „Dezvoltarea capacităților de a recepta și de a produce teste scrise și orale de diverse tipuri”.

Obiectivul de referință

1.5. „Să combine elemente ficționale și nonficționale producând texte de graniță”.

Obiectivul de evaluare

Elevul/candidatul trebuie să fie capabil să redacteze, în 1-2 pagini, o compunere de tip scrisoare familială, respectând toate categoriile de norme cunoscute (de redactare, de ortografie, de punctuație etc.).

Exemplul anterior subliniază necesitatea corelării permanente a obiectivelor de evaluare formulate de către profesori în activitatea lor curentă, sau de către cei care proiectează instrumentele de evaluare, cu celelalte categorii de obiective aflate în documentele oficiale.

Sumar de idei

Concepte-cheie:

- reforma curriculară;
- relația curriculum-evaluare;
- efectul de *feedback*;
- efectul de *backwash*;
- standarde de performanță;
- obiective cadru/obiective de referință/obiective de evaluare;
- operaționalizare;
- taxonomii ale obiectivelor educaționale.

Idei principale:

- Evaluarea educațională este concepută ca o parte activă, integrantă a curriculum-ului;
- *Feedback-ul* formativ, ca și cel sumativ, reprezintă fațete ale unor efecte ce pot și trebuie să fie monitorizate la nivelul procesului de evaluare continuă;
- Efectul de *backwash* al evaluării/examinării asupra curriculum-ului poate fi ponderat prin armonizarea relației dintre evaluarea curentă și examinare;
- Standardele de performanță reprezintă, din perspectiva evaluării și examinării, acele elemente „realizate”, „atinse”, „demonstrate” de către educabili prin intermediul performanței realizate;
- Din perspectiva evaluării, operaționalizarea obiectivelor reprezintă un element esențial al proiectării instrumentelor de evaluare/examinare.

Capitolul III. Metode și instrumente de evaluare

Etimologic, conceptul de metodă derivă din grecescul „methodos”, care semnifică „drum spre”..., „cale de urmat” pentru atingerea unui obiectiv bine precizat și definit, un demers de cercetare sau investigare a modalităților eficiente de cunoaștere a realității, de reflecție asupra ei și de transformare a acesteia pe baza experienței anterioare a elevilor. Circumscrișă acțiunii instructiv-educative, metoda devine acea cale eficientă de organizare și de conducere a învățării sub seninul comun al aceluiași mod de acțiune obiectivat în interacțiunea dintre profesor și elevii săi.

Și din perspectiva evaluativă, metoda cumulează aceleași semnificații educaționale, reprezentând un mod eficient de realizare a demersului inițiat în vederea atingerii obiectivelor de evaluare propuse.

Metoda de evaluare este o cale prin care profesorul oferă elevilor posibilitatea de a demonstra nivelul de stăpânire a cunoștințelor, de formare a diferitelor capacități testate prin utilizarea unei diversități de instrumente adecvate scopului urmărit.

Relația dintre metoda de evaluare și instrumentul de evaluare trebuie interpretată ca o relație de dependență univocă a instrumentului de metodă, în sensul că cel dintâi își subsumează valențele formative și operaționale realizării perspectivei metodologice propuse.

Astfel, metoda de evaluare, în ansamblul său, vizează întregul demers de proiectare și realizare a actului evaluativ, de la stabilirea obiectivelor de evaluare și până la construirea și aplicarea instrumentului de evaluare prin care intenționăm să obținem informațiile necesare și relevante pentru scopurile propuse. Din această perspectivă, **instrumentul de evaluare** este parte integrantă a metodei, fiind cel care concretizează la nivel de produs opțiunea metodologică a profesorului pentru testarea performanțelor elevului într-o situație educațională bine definită.

În dinamica procesului evaluativ cele două etape importante, măsurarea și aprecierea capătă semnificație și coerentă în funcție de metoda utilizată în susținerea acestui demers. În contextul metodologic descris pentru un anumit tip de evaluare, instrumentul în sine devine partea operațională relaționată cu sarcina de lucru prin intermediul căreia elevul demonstrează abilități și capacități specifice situației de învățare și/sau de evaluare.

Instrumentul de evaluare este cel care pune în valoare atât obiectivele de evaluare, cât și demersul inițiat pentru a atinge scopul propus, uneori reușind chiar o schimbare a modului de abordare a practicii evaluative curente sau a celei de examen. Utilizarea din ce în ce mai susținută a unor instrumente de evaluare care valorizează sarcinile de lucru centrate pe competențe și abilități de ordin practic ar putea revigora întreaga viziune a profesorului asupra a ceea ce trebuie evaluat, când, cum și mai ales, cu ce.

Aceasta este și perspectiva din care vom aborda relația dintre metoda și instrumentul de evaluare în acest capitol.

În lucrările de specialitate, metodologia evaluării discriminează între **metodele tradiționale de evaluare** și cele **complementare**.

3.1. Metode tradiționale de evaluare

Metodele tradiționale de evaluare au căpătat această denumire datorită consacării lor în timp ca fiind cele mai des utilizate. Din această categorie fac parte:

- probele orale;
- probele scrise;
- probele practice.

Pentru asigurarea unui consens din punct de vedere conceptual vom adopta în continuare următoarea definiție de lucru (Stoica, Musteață, 1997, p.103):

Probă = orice instrument de evaluare proiectat, administrat și corectat de către profesor.

3.1.1. Probele orale

Probele orale reprezintă metoda de evaluare cel mai des **utilizată** la clasă. Datorită fidelității și validității lor scăzute, aceste probe nu sunt recomandabile în situații de examen, ele fiind caracteristice în principal disciplinelor care presupun demonstrarea unor capacități și abilități dificil de surprins prin intermediul probelor scrise (de exemplu, capacitatea de comunicare verbală).

Avantajele utilizării probelor orale vizează:

- flexibilitatea și adecvarea individuală a modului de evaluare prin posibilitatea de a alterna tipul întrebărilor și gradul lor de dificultate în funcție de calitatea răspunsurilor oferite de către elev;
- posibilitatea de a clarifica și corecta imediat eventualele erori sau neînțelegeri ale elevului în raport de un conținut specific;
- formularea răspunsurilor urmărind logica și dinamica unui discurs oral, ceea ce oferă mai multă libertate de manifestare a originalității elevului, a capacității sale de argumentare etc.;
- nu în ultimul rând, tipul de interacțiune directă creată între evaluator și evaluat (profesor și elev), de natură să stimuleze modul de structurare a răspunsurilor de către elev, încurajând și manifestări care permit evaluarea comportamentului afectiv-atitudinal.

Este necesar, însă, să fie avute în vedere și limitele acestor probe dintre care menționăm:

- diversele circumstanțe care pot influența obiectivitatea evaluării atât din perspectiva profesorului, cât și din cea a elevului. De exemplu, gradul diferit de dificultate al întrebărilor de la un elev la altul, variația comportamentului evaluatorului etc. generează o puternică varietate interindividuală și intraindividuală între evaluatori sau la același evaluator în momente diferite, la fel cum starea emoțională a elevului în momentul răspunsului influențează performanța acestuia din punct de vedere al calității prestației sale;
- nivelul scăzut de validitate și fidelitate;
- consumul mare de timp, având în vedere că elevii sunt evaluați individual.

Decizia profesorului de a utiliza probele orale în anumite situații ale practicii școlare curente trebuie să se întemeieze pe rațiuni care țin de:

- obiectivele evaluării în situația concretă;
- tipul de evaluare promovat;
- numărul elevilor;
- timpul disponibil și resursele materiale alocate;
- tipul de informație pe care profesorul dorește să o obțină prin răspunsurile elevilor;
- natura și specificul disciplinei.

3.1.2. Probele scrise

Probele scrise sunt practicate, și uneori chiar preferate, datorită unora dintre avantajele lor imposibil de ignorat în condițiile în care se dorește eficientizarea procesului de instruire și creșterea gradului de obiectivitate în apreciere. Dintre acestea menționăm:

- economia de timp pe care o realizează în cadrul bugetului alocat relației predare-învățare-evaluare. Probele scrise permit evaluarea unui număr mare de elevi într-un timp relativ scurt.
- acoperirea unitară ca volum și profunzime pe care acest tip de probe o asigură la nivelul conținutului evaluat. Probele scrise fac posibilă evaluarea tuturor elevilor asupra aceleiași secvențe curriculare; ceea ce face comparabile rezultatele elevilor, iar evaluarea în sine mai obiectivă.
- posibilitatea evaluatorului de a emite judecăți de valoare mult mai obiective, întemeiate pe existența unor criterii de evaluare clar specificate și prestabilite.
- posibilitatea elevilor de a-și elabora răspunsul în mod independent, reflectând cunoștințe și capacități demonstrate într-un ritm propriu.
- diminuarea stărilor tensionale, de stress, care pot avea un impact negativ asupra performanței elevilor timizi sau cu alte probleme emoționale.

Dezavantajul major este presupus de relativa întârziere în timp a momentului în care se realizează corectarea unor greșeli sau completarea unor lacune identificate.

3.1.3. Probele practice

Probele practice sunt utilizate în vederea evaluării capacității elevilor de a aplica anumite cunoștințe teoretice, precum și a nivelului de stăpânire a priceperilor și deprinderilor de ordin practic. Cu toate că activitățile practice oferă posibilitatea elevului de a-și dezvolta atât competențele generale (comunicare, analiză, sinteză, evaluare), cât și pe cele specifice, aplicative (utilizarea datelor, a

instrumentelor de lucru, interpretarea rezultatelor), evaluarea elevilor prin probe practice, atât în situații de examinare curentă, cât și în situații de examen, este foarte puțin pusă în valoare.

Pentru realizarea cu succes a unei activități practice, este normal ca încă de la începutul anului școlar, elevii să fie avizați asupra (Neacșu, Stoica - coord., 1996, p. 76):

- tematicii lucrărilor practice;
- modului în care ele vor fi evaluate (baremele de notare);
- condițiilor care le sunt oferite pentru a realiza aceste activități (aparate, unelte, săli de sport etc.).

Un tip specific de probă practică îl constituie **activitățile experimentale** în contextul disciplinelor cu caracter practic-aplicativ (vezi și Neacșu, Stoica - coord., 1996, pp. 62-70). Pentru a putea facilita sarcina elevului de a-și însuși, pe de o parte, cunoștințe de bază și, pe de altă parte, o metodologie de investigație, alături de aspectul teoretic al activității de învățare trebuie să fie prezentă și activitatea practică experimentală. În cadrul acestor activități, elevul învață și descoperă mecanisme specifice de investigație, de observare și experimentare, de reflecție etc.

Prin intermediul acestor activități experimentale, care reprezintă în primul rând situații de învățare, profesorul își propune să evalueze capacități variate ale elevilor, care nu pot fi surprinse prin intermediul altor tipuri de probe. Dintre acestea menționăm, spre exemplu:

- capacitatea de a manipula corect aparatura, substanțele;
- capacitatea de a utiliza aparatele de măsură;
- capacitatea de a utiliza „limbajele specifice” (coduri, instrumente matematice);
- capacitatea de a înregistra și de a prezenta cu claritate datele și rezultatele obținute etc.

Caracteristica principală a activităților experimentale este reprezentată de **caracterul preponderent formativ** al acestora, manifestat în domenii de activitate variate, cum ar fi:

- verificarea fenomenelor, legilor, relațiilor cunoscute;
- planificarea și realizarea unui experiment;
- determinarea valorilor și mărimilor;
- observarea și studierea unor fenomene determinate etc.

Modalitățile de evaluare a activităților experimentale variază în general între două extreme:

(i) se realizează experimente scurte, urmate de întrebări la care elevii răspund în scris sau oral, apoi se notează datele experimentale și se interpretează rezultatele.

(ii) se pornește de la o problemă bine aleasă, clară, ce urmează a fi studiată în vederea determinării unei legi, a studiului unui anumit sistem.

Și într-un caz și în celălalt, procesul de evaluare vizează elemente specifice care trebuie avute în vedere încă din momentul în care se proiectează activitatea experimentală.

3.2. Metode complementare de evaluare

Așa cum menționam anterior, metodele tradiționale de evaluare, concepute ca realizând un echilibru între probele orale, scrise și cele practice, constituie la momentul actual elementele principale și dominante în desfășurarea actului evaluativ. Pornind de la această realitate obiectivă, strategiile moderne de evaluare caută să accentueze acea dimensiune a acțiunii evaluative care să ofere elevilor suficiente și variate posibilități de a demonstra ceea ce știu (ca ansamblu de cunoștințe), dar, mai ales, ceea ce pot să facă (priceperi, deprinderi, abilități).

Preocuparea continuă a practicienilor din domeniul educației de a găsi și valorifica noi tehnici și proceduri de evaluare, mai ales pentru măsurarea acelor obiective aparținând domeniului afectiv, mai greu cuantificabile prin metodele clasice de evaluare, s-a concretizat în identificarea și utilizarea unor metode de evaluare care pot reprezenta o alternativă viabilă la formulele evaluative tradiționale, fiind complementare acestora.

Principalele **metode complementare de evaluare**, al căror potențial formativ susține individualizarea actului educațional prin sprijinul acordat elevului sunt:

- observarea sistematică a activității și a comportamentului elevilor;

- investigația;
- proiectul;
- portofoliul;
- autoevaluarea.

3.2.1. Observarea sistematică a activității și a comportamentului elevilor

Observarea sistematică a comportamentului elevilor în timpul activității didactice este o tehnică de evaluare care furnizează profesorului o serie de informații utile, diverse și complete, greu de obținut altfel prin intermediul metodelor de evaluare tradiționale (Stoica, Musteață, 1997, p. 109).

În sensul său cel mai general, *observația constă în investigarea sistematică, pe baza unui plan dinainte elaborat și cu ajutorul unor instrumente adecvate, a acțiunilor și interacțiunilor, a evenimentelor, a relațiilor și a proceselor dintr-un câmp social dat.*

Conform acestei accepțiuni, profesorul care-și propune să utilizeze această metodă pentru a obține informații relevante asupra activității elevilor săi, din perspectiva capacității lor de acțiune, relaționare, a competențelor și abilităților de care dispun, trebuie să utilizeze un instrumentar adecvat obiectului observării. Pentru a înregistra aceste informații, profesorul are la dispoziție în mod practic trei modalități (vezi și Gronlund, 1981, pp. 434-450):

- **fișa de evaluare (calitativă);**
- **scara de clasificare;**
- **lista de control/verificare.**

Aceste instrumente se utilizează atât pentru evaluarea procesului, cât și a produselor realizate de către elevi, cu ajutorul lor putând fi surprinse atât obiectivări comportamentale ale domeniului cognitiv, cât și ale domeniilor afectiv și psihomotor.

Acest tip de observație *participativă* practică de profesor, utilizând cele trei modalități menționate, este, în esență, subiectivă, dar poate să-și sporească gradul de obiectivitate dacă se concentrează atenția asupra modului de elaborare și utilizare a instrumentelor. În ceea ce privește *costurile* implicate de această metodă, ea este *ieftină, dar mare consumatoare de timp*.

• **Fișa de evaluare** este completată de către profesor, în ea înregistrându-se date factuale despre evenimentele cele mai importante pe care profesorul le identifică în comportamentul sau în modul de acțiune al elevilor săi (fapte remarcabile, probleme comportamentale, evidențierea unor aptitudini deosebite într-un domeniu sau altul etc.). La acestea se adaugă interpretările profesorului asupra celor întâmplate, permițându-i acestuia să surprindă modelul comportamental al elevilor săi.

Deci, profesorul, într-o primă instanță, este cel care decide ce comportament va fi observat și înregistrat, configurând astfel aria comportamentală, eventualele limite ale observațiilor înregistrate, precum și gradul de extindere în utilizare al acestor observații.

Exemplu (adaptat după Gronlund, 1981, p. 435):

<p>Clasa.....Numele elevului</p> <p>Disciplina..... Data.....</p> <p>Eveniment</p> <p>Pentru ora respectivă elevii au avut de pregătit o compunere în care să descrie cea mai dragă persoană. Florin a fost singurul elev care a declarat că nu și-a făcut tema. La jumătatea orei, colegul lui de bancă mi-a spus că Florin scrisese de fapt tema, dar nu era o compunere, ci o poezie despre mama sa, pe care însă nu vrea să o citească, de teamă că ceilalți colegi vor râde de el.</p> <p>La sfârșitul orei, Florin m-a întrebat dacă poate citi poezia scrisă de el. A citit poezia cu o voce joasă, uitându-se tot timpul în hârtie, mișcându-și piciorul drept înainte și înapoi și trăgându-se de gulerul cămășii. Când a terminat, colegul lui din ultima bancă a spus: „N-am auzit nimic. Poți să mai citești poezia încă o dată, mai tare?”. Florin a spus: „Nu” și s-a așezat.</p>
--

Interpretare:

Florin dovedește un talent literar real pentru vârsta lui, dar nu are încă încredere în forțele sale. Nu îi place să citească în fața clasei și îi este teamă de reacția celorlalți. Refuzul de a citi încă o dată poezia se datorează probabil timidității sale și faptului că nimeni nu l-a încurajat până acum să facă acest lucru. Talentul său trebuie stimulat, ca și încrederea de sine.

Un **avantaj** important al acestor fișe de evaluare este acela că nu depind de capacitatea de comunicare a elevului cu profesorul, profesorul fiind cel care înregistrează și interpretează comportamentul tipic sau alte produse și performanțe ale elevului.

Un **dezavantaj** de care trebuie ținut seamă este acela al marelui *consum de timp* pe care îl implică (timpul necesar profesorului, spre exemplu, pentru a nota descrierile verbale ale comportamentului elevilor săi), la care se adaugă faptul că *aceste observații nu au o cată ridicată de obiectivitate*, ceea ce are repercursiuni asupra fidelității acestor înregistrări.

• **Scara de clasificare** însumează un set de caracteristici (comportamente) ce trebuie supuse evaluării, însoțit de un anumit tip de scară, de obicei scara *Likert*. Potrivit acestui tip de scară, elevului îi sunt prezentate un număr de enunțuri în raport de care acesta trebuie să-și manifeste acordul sau dezacordul, discriminând între 5 trepte: **puternic acord; acord; indecis** (neutru); **dezacord; puternic dezacord**.

Exemplu:

- | | |
|--|--|
| 1. Particip cu plăcere la activitățile organizate pe grupuri de lucru. | puternic dezacord dezacord neutru acord puternic acord |
| 2. Îmi place ca părerile pe care le exprim să fie întotdeauna respectate de ceilalți | puternic dezacord dezacord neutru acord puternic acord |

Secretul construirii unei astfel de scări îl constituie *redactarea unui bun enunț* la care elevul să poată răspunde. În continuare prezentăm câteva **sugestii** pentru construirea unei scări Likert.

• Fiecare enunț să cuprindă cuvinte familiare elevului, utilizând limbajul acestuia și nu unul puternic specializat sau tehnic. Deci, *enunțuri cu o structură simplă*.

• Fiecare enunț să fie exprimat *clar pozitiv* sau *clar negativ* (să reprezinte poziții clar pozitive sau negative). Este bine ca enunțurile să fie experimentate înainte de introducerea lor în cadrul instrumentului.

• Lista finală de enunțuri trebuie să conțină *un număr aproximativ egal de enunțuri pozitive și negative*.

• *Fiecare enunț trebuie să producă informația necesară*. Nu se introduce un enunț doar pentru că „este interesant de văzut ce răspund elevii la el”. Enunțul trebuie să facă referire clară la atitudinea sau opinia despre care dorim să aflăm informații.

• **Lista de control/verificare**, deși pare asemănătoare cu scara de clasificare ca manieră de structurare (un set de enunțuri, caracteristici, comportamente etc.), se deosebește de aceasta prin faptul că prin intermediul ei doar se constată *prezența* sau *absența* unei caracteristici, comportament etc., fără a emite o judecată de valoare oricât de simplă.

Exemplu (desfășurarea unei activități experimentale în laborator):

Elevul:	
- a urmat instrucțiunile specifice activității.	Da Nu
- a cerut ajutor atunci când a avut nevoie.	Da Nu
- a colaborat cu ceilalți colegi pentru a realiza produsul final.	Da Nu
- a finalizat sarcina de lucru.	Da Nu
- a făcut curat pe masa de lucru.	Da Nu

Un **avantaj** al listei de control este acela că se *elaborează relativ ușor, fiind și simplu de aplicat elevilor*.

3.2.2. Investigația

Investigația oferă posibilitatea elevului de a aplica în mod creativ cunoștințele însușite, în situații noi și variate, pe parcursul unei ore sau unei succesiuni de ore de curs. Această metodă presupune definirea unei sarcini de lucru cu instrucțiuni precise, înțelegerea acesteia de către elevi înainte de a trece la rezolvarea propriu-zisă prin care elevul demonstrează, și exersează totodată, o gamă largă de cunoștințe și capacități în contexte variate.

Investigația oferă, de asemenea, posibilitatea elevului de a se implica activ în procesul de învățare, realizând permanente integrări și restructurări în sistemul noțional propriu, ceea ce conferă cunoștințelor un caracter operațional accentuat. Investigația stimulează inițiativa elevilor pentru luarea deciziilor, oferind un nivel de înțelegere mult mai profundă asupra evenimentelor și fenomenelor studiate, motivând în același timp elevii în realizarea activităților propuse.

Prin realizarea unei investigații pot fi urmărite ca elemente esențiale:

- înțelegerea și clarificarea sarcinii de lucru;
- identificarea procedeeleor pentru obținerea informațiilor necesare;
- colectarea și organizarea datelor sau informațiilor necesare;
- formularea și testarea unor ipoteze de lucru;
- schimbarea planului de lucru sau a metodologiei de colectare a datelor, dacă este necesar;
- colectarea altor date, dacă este necesar;
- motivarea opțiunii pentru anumite metode folosite în investigație;
- sciirea/prezentarea unui scurt raport privind rezultatele investigației.

Demersul investigativ poate fi raportat la trei etape esențiale care trebuie parcurse (vezi și Fairbrother și alții, 1992, p. 12):

- definirea problemei,
- alegerea metodei/metodologiei adecvate,
- identificarea soluțiilor.

Abordarea investigației se poate plasa la diferite niveluri de deschidere, în funcție de abilitățile și competențele pe care le solicită din partea elevilor.

În acest sens, se pot compara, spre **exemplu**, următoarele tipuri de investigații:

- A. Investigați cum lumina influențează procesul de fotosinteză.
- B. Studiați factorii care influențează procesul de fotosinteză.

Investigația B este complexă în comparație cu investigația A deoarece nu are specificate nici un fel de variabile prin modul de formulare a sarcinii de lucru. În această situație, deschiderea pe care o oferă în proiectarea demersului investigativ creează posibilități elevului de a-și demonstra deprinderi și competențe superioare celor presupuse de realizarea investigației A, care reprezintă un tip de sarcină de lucru mult mai structurată.

Din acest punct de vedere, etapele enunțate anterior, specifice unui demers investigativ pot fi descrise astfel:

- definirea problemei poate fi făcută:

(i) într-un mod prescriptiv, care direcționează clar activitatea elevului. În acest caz, variabilele implicate în investigație sunt specificate și operaționalizate adecvat. Demersul propus este bine structurat.

SAU

(ii) într-un mod care accentuează dimensiunea de explorare. În acest caz, variabilele nu sunt specificate, ci doar domeniul de investigație.

- alegerea metodei/metodologiei adecvate poate fi realizată:

(i) fie de către profesor, care arată elevilor ce să facă, fie furnizează toate informațiile necesare legate de aparatura adecvată, de instrumente etc.

(ii) fie de către elev, care are totală libertate de a alege metodele corespunzătoare demersului inițiat.

- identificarea soluțiilor poate presupune formularea:

(i) unei singure soluții acceptabile;

SAU

(ii) mai multor soluții acceptabile.

Cele două puncte extreme prezentate în raport cu etapele menționate anterior descriu un continuum care evoluează pe niveluri diferite de deschidere de la activități investigative cu un caracter bine structurat către cele cu un caracter de explorare mult accentuat. Gradul de libertate pe care îl poate oferi o investigație depinde în același timp și de nevoile, interesele manifestate de către elev, dar și de experiențele sale anterioare. Profesorul poate adapta acest demers pentru diferite categorii de elevi, oferindu-le acestora fie mai multe informații, instrucțiuni clare de urmat, fie lăsându-le totală libertate de inițiativă în realizarea activității.

Sarcinile de lucru adresate elevilor de către profesor în realizarea unei investigații, pot varia ca nivel de complexitate a cunoștințelor și competențelor implicate, după cum urmează:

- **simpla descriere a caracteristicilor unor obiecte**, lucruri desprinse din realitatea imediată sau fenomene observate direct de către elev și comunicarea în diferite moduri a observațiilor înregistrate, prin intermediul desenelor, graficelor, tabelelor sau hărților;

- **utilizarea unor echipamente simple** pentru a face observații, teste referitoare la fenomenele supuse atenției elevilor. Aceste observații constituie baza pentru realizarea unor comparații adecvate între fenomenele respective sau între ceea ce au înregistrat direct și ceea ce au presupus că se va întâmpla (confirmarea sau nu a predicțiilor făcute).

- **identificarea factorilor implicați în contextul** supus observației, prin intermediul aparaturii specifice. Elevii pot repeta observațiile și măsurătorile pentru a oferi explicații pertinente diferențelor sesizate în derularea activității. Pe baza înregistrării sistematice a observațiilor și rezultatelor măsurătorilor se emit concluzii prezentate într-o formă științifică și argumentate logic pentru confirmarea predicțiilor formulate.

Selectarea echipamentului adecvat realizării sarcinii, efectuarea unor serii de măsurători, înregistrarea observațiilor specifice, prezentarea acestora sub formă de concluzii, utilizând tabele, grafice și hărți sunt tot atâtea operații care antrenează elevii într-o formă de activitate teoretico-practică cu puternice **valențe formative**.

În momentul în care profesorul propune elevilor săi realizarea unei investigații, trebuie mai întâi să reflecteze asupra răspunsului la următoarele întrebări:

- Ce așteptări are de la elevii săi?

- Ce cunoștințe, noțiuni aplică elevii, ce deprinderi și abilități își exersează și la ce nivel de înțelegere se va plasa întregul demers investigativ?

- Care vor fi rezultatele învățării în urma realizării acestei activități și cum vor fi ele apreciate?

Exemplu:

Care este cel mai bun material pentru realizarea unui costum de protecție într-un laborator chimic?

Acest tip de investigație poate fi utilizat pentru o varietate de scopuri la diferite niveluri de vârstă ale elevilor. Domeniul de referință la care se raportează elevii în realizarea activității este cel legat de diferitele tipuri de materiale. Pentru a răspunde cerinței, aceștia trebuie să aplice cunoștințe legate de proprietățile materialelor și de mediul lor de aplicabilitate, fiind necesar să dezvolte strategii pentru argumentarea alegerii făcute. Elevii trebuie să decidă asupra variabilelor implicate de demersul investigativ, să identifice cele mai bune proceduri metodologice, precum și strategia optimă care să conducă la răspunsul corect pe baza observațiilor înregistrate.

Activitatea didactică desfășurată prin intermediul acestei practici evaluative, poate să fie organizată individual sau pe grupuri de lucru, iar *aprecierea* modului de realizare a investigației este, de obicei, de **tip holistic**. Criteriile care se stabilesc pentru aprecierea nivelului de realizare al fiecărei etape trebuie clar precizate și cunoscute de către elevi încă de la începutul activității. Aceste criterii descriu performanța așteptată corespunzător calificativelor sau notelor/punctajelor care se alocă acestora.

Organizarea activității de lucru individual sau pe **grupuri** trebuie să țină cont de:

- natura investigației / scopul pe care aceasta și-l propune;
- integrarea investigației ca parte componentă a lecției sau ca activitate în sine;
- modul în care elevii sunt deja obișnuiți să lucreze în clasă.

Adesea, investigația presupune participarea profesorului. Am rolul de consilier/sfătuitor al elevului în momentele în care acesta solicită sprijin, întreaga libertate în elaborarea planului de lucru și Am derularea observațiilor/experimentelor specifice revenindu-i elevului. În cazul în care se lucrează în grup, profesorul trebuie să găsească timpul necesar pentru a discuta activitatea tuturor grupurilor implicate în realizarea activității respective.

În situația în care **lecția este structurată pe baza unei investigații, de regulă**, se urmărește realizarea următoarelor **etape** (vezi și Fairbrother și alții, 1992, p. 22):

(i) concentrarea asupra problemei de investigat;

(ii) explorarea posibilităților de rezolvare/realizare;

(iii) raportarea/prezentarea rezultatelor obținute;

(iv) consolidarea cunoștințelor/noțiunilor presupuse de realizarea investigației, valorificarea ideilor elevilor și a demersurilor convergente în identificarea soluțiilor;

(v) aplicarea rezultatelor obținute în sensul relațiilor interdisciplinare care pot fi stabilite și al legăturii cu viața reală.

În acest caz, investigația devine element de bază în sprijinirea demersului de învățare prin descoperire. Cum, însă, predarea, învățarea și evaluarea sunt inter-dependente în contextul strategiilor educaționale modeme, investigația reprezintă prin valențele sale formative indiscutabile o metodă care deservește toate cele trei activități menționate anterior.

Prin intermediul investigației, profesorul poate urmări **procesul** sau/și **realizarea unui produs** sau/și **atitudinea** elevului. Printre **calitățile personale ale elevului** pe care investigația le pune în valoare se pot număra:

- creativitatea și inițiativa;
- cooperarea și participarea la lucrul în echipă;
- preluarea conducerii/inițiativei în cadrul grupului;
- constanța și concentrarea atenției;
- perseverența;
- flexibilitatea gândirii și deschiderea către noi idei.

Investigația, ca instrument de evaluare, constituie o reală șansă pentru elev de a-și pune în valoare potențialul creativ în aplicarea cunoștințelor asimilate, în explorarea situațiilor noi (învățare euristică) sau foarte puțin asemănătoare în raport cu situația anterioară. În același timp, acest tip de activitate dezvoltă capacitatea de argumentare, de gândire logică, de rezolvare a problemelor etc.

3.2.3. Proiectul

Proiectul presupune un demers evaluativ **mult mai amplu** decât investigația. Proiectul începe în clasă, prin definirea și înțelegerea sarcinii de lucru, eventual și prin începerea rezolvării acesteia - se continuă acasă pe parcursul a câtorva zile sau săptămâni, timp în care elevul are permanente consultări cu profesorul și se încheie tot în clasă, prin prezentarea în fața colegilor a unui raport asupra rezultatelor obținute și, dacă este cazul, a produsului realizat. Ca și investigația, proiectul are mai multe etape și poate fi realizat individual sau în grup.

Etapile proiectului presupun direcționarea eforturilor elevilor în două direcții la fel de importante din punct de vedere metodologic și practic: **colectarea datelor** și **realizarea produsului**. Fiecare direcție conține elemente care conferă specificitate proiectului în funcție de disciplina pe care o vizează.

Printre **capacitățile elevilor** posibil de evaluat prin intermediul acestei metode se pot enumera (vezi și Stoica, Musteață, 1997, p.115):

- adecvarea metodelor de lucru și a instrumentarului ales la obiectivele pro-puse prin proiect;
- folosirea corespunzătoare a materialelor și a echipamentelor din dotare;
- oferirea unei soluții corecte (rezolvarea de probleme);
- realizarea cu acuratețe a produsului, din punct de vedere tehnic;
- posibilitatea generalizării problemei/soluției;
- prezentarea proiectului.

Proiectul ca instrument de evaluare poate lua forma unei sarcini de lucru **individuale** sau **de grup**, ținând cont și de faptul că o bună parte a activității presupuse de acesta poate fi realizată și în afara orelor de curs.

Alegerea temei pentru proiect poate fi făcută de către profesor sau poate aparține elevului însuși. Indiferent cine este cel care decide asupra tematicii proiectului, există câteva **întrebări** care țin de dimensiunea evaluativă și cărora trebuie să li se răspundă înainte de realizarea proiectului în sine (vezi și Stoica, Musteață, 1997, p. 116).

- Pe ce anume se va centra demersul evaluativ? Pe proces, pe produsul final sau vor fi luate în considerație ambele perspective? Este vreuna dintre ele mai importantă din punct de vedere al evaluării în contextul particular dat?
- Care va fi rolul profesorului? Va avea acesta un rol de consilier permanent al elevului pe tot parcursul realizării proiectului, va fi doar evaluatorul final sau coordonatorul întregii activități?
- Care este statutul resurselor implicate în derularea proiectului? Sunt puse la dispoziția elevului de la începutul activității sau trebuie să fie identificate de către acesta și utilizate adecvat? Resursele sunt comune pentru toți elevii în realizarea aceluiași proiect etc.?
- Există o anumită structură a proiectului propusă/impusă de către profesor? Este aceasta o opțiune de organizare și structurare care aparține în exclusivitate elevului? Există anumite caracteristici ale produsului final al proiectului pe care trebuie să le îndeplinească toți elevii?

În demersul de realizare a unui proiect următorii **pași** sunt foarte important de urmărit:

- stabilirea domeniului de interes;
- stabilirea premiselor inițiale — cadru conceptual, metodologic, datele generale ale investigației/anchetei;
- identificarea și selectarea resurselor materiale;
- precizarea elementelor de conținut ale proiectului.

Exemplul pe care îl vom prezenta în continuare ca modalitate de obiectivare a acestor pași în realizarea unui proiect va avea ca domeniu de referință disciplinele socio-umane.

Raportat la primul aspect invocat anterior și care viza **clarificarea ariei de interes**, primul element care trebuie avut în vedere de către elev este acela de a aprecia **relevanța scopului/obiectivelor** pe care și le propune proiectul din perspectiva adecvării la specificul ariei curriculare/disciplinei respective. Abia după aceea, atenția poate fi concentrată asupra unui aspect particular din contextul domeniului de interes.

Spre **exemplu**, poate fi propusă o temă generală circumscrisă ariei tematice a disciplinelor socio-umane, sub denumirea de „*Imaginea de sine și reprezentările reciproce dintre indivizi în societate*”. Această temă poate include ca elemente de studiu următoarele:

- „Atitudinea tinerei generații față de valorile promovate de societatea contemporană”;
- „Conflictul dintre generații”;
- „Problema implicării și a responsabilității în soluționarea problemelor sociale” etc.

În pasul următor, când există deja conturată aria de interes și tema specifică de studiu, autorul proiectului trebuie să gândească asupra **tipului de informații** asupra căruia dorește să-și centreze analiza. Cel mai simplu mod de a rezolva această problemă este de a stabili un set de întrebări esențiale, care pot deveni în același timp și elemente-cheie în jurul cărora se poate centra conținutul proiectului. Răspunsurile la aceste întrebări sprijină elevul în clarificarea cadrului general de acțiune, precum și a perspectivei de abordare a temei stabilite.

Informația utilă pentru **identificarea și selectarea resurselor materiale** se poate colecta fie din **surse secundare** (lucrări de specialitate, articole, alte materiale/documente scrise), fie din **surse primare** (informația obținută în urma observării sistematice, a interviurilor, a anchetelor etc.).

Elementele de conținut ale proiectului se pot organiza după următoarea structură:

(i) Pagina de titlu pe care, de obicei, se consemnează tema proiectului, numele autorului, școala, perioada în care s-a elaborat proiectul;

(ii) Cuprinsul proiectului care prezintă titlurile capitolelor și subcapitolelor pe care se structurează lucrarea;

(iii) Introducerea care include prezentarea cadrului conceptual și metodologic căruia i se circumscrie studiul temei propuse;

(iv) Dezvoltarea elementelor de conținut, a capitolelor și subcapitolelor care oferă substanță și fundament analizei inițiate;

(v) Concluziile care sintetizează elementele de referință desprinse în urma studiului temei respective, sugestii/propuneri de ameliorare a aspectelor vulnerabile semnalate;

(vi) Bibliografia;

(vii) Anexa care poate include toate materialele importante rezultate în urma aplicării unor instrumente de investigație (grafice, tabele, chestionare, fișe de observație etc.) și care susțin demersul inițiat.

Pentru realizarea unei evaluări cât mai obiective a proiectului trebuie avute în vedere câteva **criterii generale de evaluare**, criterii care țin de **aprecierea calității proiectului** (sau de calitatea produsului), pe de o parte, și altele care țin de **calitatea activității elevului** (sau de calitatea procesului), pe de altă parte. Fiecare dintre cele două categorii de criterii obiectivează aspecte concrete care vizează modul de realizare și prezentare a unui proiect și care vor fi menționate în continuare.

Criterii generale de evaluare a proiectului

Criterii care vizează calitatea proiectului (calitatea produsului)

1. Validitatea proiectului sau măsura în care acesta, prin modul de concepere și dezvoltare, se adecvează temei abordate de către autor. Se urmărește nivelul la care proiectul reușește să acopere unitar și coerent, logic și argumentat, câmpul tematic pe care și l-a propus să-l exploreze.

2. Completitudinea proiectului reflectată în completitudinea modului de abordare a temei ce face obiectul studiului său. Se urmărește modul în care au fost valorizate în cadrul proiectului conexiunile și perspectivele interdisciplinare, competențele și abilitățile de ordin teoretic și practic și maniera în care acestea se regăsesc în conținutul științific.

3. Elaborarea și structurarea proiectului, criteriu care vizează evidențierea competențelor demonstrate de către elev în elaborarea proiectului (acuratețea și rigurozitatea demersului științific, logica și argumentarea ideilor, coerența internă, corectitudinea ipotezelor și concluziilor etc.).

4. Calitatea materialului utilizat de către elev în realizarea proiectului se raportează direct la relevanța conținutului științific, la semnificația și acuratețea datelor colectate, precum și la opțiunea motivată pentru o strategie sau alta de prelucrare și analiză a datelor.

5. Creativitatea, ca și criteriu general în evaluarea calității proiectului, vizează gradul de noutate pe care îl aduce proiectul respectiv în abordarea tematicii propuse sau în soluționarea problemelor generate pe parcursul elaborării proiectului. Prin intermediul acestui criteriu se urmărește măsura în care elevul optează pentru o strategie de lucru clasică, tradițională în finalizarea demersului său științific, sau pentru o strategie inedită, originală, inovativă.

Criterii care vizează calitatea activității elevului (calitatea procesului)

1. Raportarea elevului la tema proiectului. Acest criteriu se referă la modul în care elevul a răspuns prin structurarea și conținutul proiectului său, cadru-lui tematic în care acesta se circumscrie. Întrebarea la care trebuie să răspundă profesorul evaluând proiectul respectiv este dacă și în ce măsură elevul a înțeles sarcina pe care a avut-o de realizat, dacă și în ce măsură a făcut ceea ce trebuie sub aspectul validității și relevanței abordării temei proiectului.

2. Performarea sarcinilor vizează, în fapt, nivelul de performanță la care se plasează elevul în realizarea diferitelor părți componente ale proiectului. Acest criteriu se poate corela foarte bine cu cel referitor la elaborarea și structurarea proiectului, având în centrul atenției valorizarea competențelor și abilităților elevului implicate în realizarea proiectului.

3. Documentarea constituie un alt criteriu esențial pentru evaluarea activității elevului în elaborarea proiectului. Există și aici câteva elemente care trebuie luate în considerare. Documentarea a fost realizată în totalitate de către elev sau acesta a fost sprijinit permanent de către profesor? Documentarea presu-pune, printre altele și identificarea bibliografiei necesare. Această bibliografie a fost identificată de către elev în mod independent sau a fost furnizată de către profesor? Bibliografia aferentă proiectului se adecvează temei, este relevantă pentru câmpul tematic respectiv? Include elemente de referință? A fost prelucrată corespunzător printr-un efort de elaborare și regândire a ideilor sau este o simplă compilație care se regăsește în conținutul proiectului? Sunt numai câteva întrebări asupra cărora profesorul trebuie să reflecteze în evaluarea proiectului prin raportare la criteriul menționat.

4. Nivelul de elaborare și comunicare reprezintă unul dintre criteriile în funcție de care se realizează analiza de conținut a proiectului. Nivelul de comunicare pe care îl evidențiază elevul în momentul prezentării proiectului poate să se situeze, fie într-un plan empiric, factual, ceea ce presupune o simplă enumerare a problemelor vizate de conținutul tematic, fie într-un plan analitic, al explicației argumentate sau într-unul de ordin evaluativ care apelează la judecăți de valoare și aprecieri personale în susținerea punctelor de vedere și a opiniilor. Forța comunicării este sporită de utilizarea în prezentarea proiectului a desenelor sau a altor elemente de grafică, care pot accesibiliza informația oferită, precum și de acuratețea realizării calculelor în anumite etape ale proiectului.

5. Greșelile sau, mai bine zis, tipul acestora poate constitui un criteriu semnificativ în aprecierea proiectului. Pornind de la nivelul la care se pot manifesta greșelile (conținut științific sau prezentare) și până la identificarea practică a tipurilor frecvente de greșeli apărute (scuzabile, acceptabile sau fundamentale), analiza acestor greșeli poate oferi suficiente argumente profesorului în întregi-rea imaginii pe care și-o construiește despre proiect și realizarea sa în ansamblu.

6. Creativitatea se manifestă și sub aspectul calității activității elevului din punct de vedere personal, reflectându-se în produse originale care poartă amprenta efortului și/sau „inspirației” autorului de proiect. Originalitatea stilului său, pusă în valoare în etapele realizării proiectului sau în momentul prezentării acestuia, conduce la aprecierea în consecință a produsului final: proiectul.

7. Calitatea rezultatelor, din perspectiva valorificării proiectului în sfera practică a activității, conferă un anumit grad de utilitate proiectului în sine și activității desfășurate de către cel care l-a realizat. Aplicabilitatea rezultatelor asigură proiectului confirmarea practică a ideilor și a strategiei utilizate în elaborarea sa.

În termeni mult mai generali, criteriile evocate anterior ca model pentru evaluarea unui proiect pot fi structurate și altfel, sub forma unor criterii mult mai sintetice, cu un nivel de acoperire mult mai mare.

Exemplu de structurare a criteriilor de evaluare a unui proiect:

- Stabilirea scopului/obiectivelor proiectului și structurarea conținutului;
- Activitatea individuală realizată de către elev (investigație, experiment, anchetă etc.);
- Rezultate, concluzii, observații. Aprecierea succesului proiectului în termeni de eficiență, validitate, aplicabilitate etc.
- Prezentarea proiectului (calitatea comunicării, claritate, coerență, capacitate de sinteză etc.);
- Relevanța proiectului (utilitate, conexiuni interdisciplinare etc.).

Opțiunea pentru modul de definire a criteriilor de evaluare a unui proiect aparține în ultimă instanță profesorului, în funcție de nivelul de generalitate la care acesta dorește să-și plaseze demersul evaluativ.

Strategia de evaluare a proiectului, care este una de **tip holistic**, trebuie la rândul ei să fie clar definită prin criterii negociate sau nu cu elevii, astfel încât să valorizeze efortul exclusiv al elevului în realizarea proiectului.

3.2.4. Portofoliul

Utilizarea **portofoliului** ca metodă complementară de evaluare se impune din ce în ce mai mult atenției și interesului profesorilor în practica școlară curentă. Termenul este relativ nou în terminologia științelor educației, dar ideea exprimată de concept în sine nu este în mod necesar inovativă în sfera practicilor evaluative. Deși în sensul său de bază portofoliul s-a lansat în domeniul artei, și în contextul științelor educației s-a impus nevoia existenței unei metode de evaluare flexibile, complexe, integratoare, ca alternativă viabilă la modalitățile tradiționale de evaluare.

Un astfel de rol își poate asuma portofoliul în cadrul procesului de evaluare a performanțelor școlare ale elevilor. El include rezultatele relevante obținute prin celelalte metode și tehnici de evaluare (probe orale, scrise, practice, observare sistematică a activității și comportamentului elevului, proiect, autoevaluare), precum și prin sarcini specifice fiecărei discipline.

Portofoliul reprezintă „**cartea de vizită**” a elevului, urmărindu-i progresul de la un semestru la altul, de la un an școlar la altul și chiar de la un ciclu de învățământ la altul.

O funcție importantă pe care o preia portofoliul este aceea de investigare a majorității „produselor” elevilor, care, de obicei, rămân neinvestigate în actul evaluativ, reprezentând în același timp un stimulent pentru desfășurarea întregii game de activități didactice (nu doar pregătirea stereotipă pentru teste de cunoștințe). În același timp, sarcina evaluării continue este preluată cu succes și fără tensiunea pe care ar putea-o genera unele dintre metodele tradiționale de evaluare aplicate frecvent.

Prin complexitatea și bogăția informației pe care o furnizează, sintetizând activitatea elevului de-a lungul timpului (un semestru, an școlar sau ciclu de învățământ), portofoliul poate constitui parte integrantă a unei evaluări sumative sau a unei examinări.

Important rămâne **scopul** pentru care este proiectat portofoliul, ceea ce va determina și structura sa. Alături de scop, în definirea unui portofoliu, sunt la fel de relevante **contextul** și **modul de proiectare** a portofoliului.

În sens general, **scopul** unui portofoliu este acela de a confirma faptul că ceea ce este cuprins în obiectivele învățării reprezintă, în fapt, și ceea ce știu elevii sau sunt capabili să facă. În determinarea scopului unui portofoliu, profesorul trebuie să răspundă mai întâi unor întrebări de tipul:

- care este conținutul — fapte, legi, teorii etc. — asimilat în acest capitol (modul, temă, domeniu de pregătire)?
- ce ar trebui elevii să fie capabili să facă (înregistrarea unor observații, crearea unei situații-problemă și rezolvarea ei, structurarea unei argumentații sau comunicarea interpersonală în cadrul relațiilor elev-elev și profesor-elev etc.)?
- care sunt atitudinile pe care elevii ar trebui să le dezvolte în realizarea portofoliului lor?

Astfel, determinarea scopului portofoliului va suferi influența simultană a curriculum-ului și a instruirii.

Scopul portofoliului este stabilit și în funcție de **destinația** sau **destinatarul său** (persoană, instituție, comunitate etc.), având în vedere că pe baza lui se va emite o judecată de valoare asupra elevului în cauză. Astfel, dacă portofoliul va servi ca instrument de evaluare destinat profesorului (un model de portofoliu cumulativ ca sursă de informații) sau ca instrument de autoevaluare pentru elev, el poate cuprinde momentele relevante ale progresului elevului, în timp ce pentru un portofoliu care trebuie să demonstreze părinților sau comunității ceea ce elevul știe sau este capabil să facă, modelul mai adecvat este acela al selectării celor mai bune produse sau a celor mai bine realizate activități ale elevului. În consecință, utilizările portofoliului sunt numeroase, în directă relație cu scopul pentru care a fost proiectat.

Un alt element esențial al portofoliului de care trebuie să se țină seamă în elaborarea acestuia este **contextul**. Dimensiuni ale acestui concept pot fi:

- vârsta elevilor;
- specificul disciplinei;
- nevoile, abilitățile și interesele elevilor etc.

Toate aceste variabile pot induce diferențe semnificative care personalizează modul de concepere și realizare a portofoliului.

Proiectarea portofoliului include, în fapt, atât scopul, cât și contextul, elemente al căror rol a fost deja menționat. Probabil cea mai importantă decizie în proiectarea portofoliului este cea care vizează *conținutul* său (identificarea elementelor reprezentative pentru activitățile desfășurate de elev) care poate fi concretizat, spre exemplu, în:

- selecții din temele pentru acasă ale elevului;
- calendarul lunar de activitate;
- notițele din clasă;
- rapoarte de laborator;
- lucrări de cercetare;
- casetă video conținând prezentări orale ale elevului, situații de învățare în grup etc.

O altă decizie importantă în cadrul proiectării este legată de :

- cât de multe astfel de "eșantioane" ale activității elevului trebuie să conțină portofoliul;
- cum să fie ele organizate;
- cine decide selecția lor(o posibilitate poate fi : profesorul descrie cerințele de conținut ale portofoliului, iar elevul selectează probele pe care le consideră reprezentative).

Este important ca întregul conținut al portofoliului să fie raportat la anumite **cerințe-standard** clar formulate în momentul proiectării și cunoscute de către elevi înainte de realizarea efectivă a acestuia.

Prin urmare, portofoliul:

- poate fi exclusiv o sarcină a profesorului, în sensul că el este cel care stabilește scopul, contextul, realizează proiectarea lui, formulează cerințele standard și selectează produsele reprezentative ale activității elevilor

sau

- poate implica și contribuția elevilor în modul în care acesta se construiește: elevii pot alege anumite instrumente de evaluare sau eșantioane din propria activitate considerate semnificative din punct de vedere al calității lor.

Din această perspectivă, **portofoliul** stimulează creativitatea, ingeniozitatea și implicarea personală a elevului în activitatea de învățare, dezvoltând motivația intrinsecă a acestuia și oferind astfel profesorului date esențiale despre personalitatea elevului ca individualitate în cadrul grupului.

Exemplu:

Un portofoliu la **limbi moderne** ar putea cuprinde;

- lucrări scrise curente;
- teste criteriale;
- răspunsuri la chestionare/interviuri;
- chestionare/interviuri elaborate personal sau în grup;
- compuneri libere;
- creații literare proprii;
- înregistrări audio/video ale elevilor (lectură de texte, prezentare a unor produse realizate);
- prezentarea unor autori (fișe de autori) sau opere literare;
- postere, colaje, machete, desene, caricaturi etc.;
- contribuții la reviste școlare.

Fiecare dintre aceste elemente constitutive ale portofoliului a fost evaluat separat de către profesor la momentul respectiv. Dacă se dorește și o apreciere globală a portofoliului, atunci profesorul va stabili **criterii** clare, **holistice** de evaluare, care vor fi comunicate elevilor înainte ca aceștia să înceapă proiectarea portofoliului. Aceste criterii pot fi elaborate în exclusivitate de către profesor sau pot fi dezvoltate în cooperare cu elevii. În această din urmă situație, elevii vor înțelege și accepta mult mai bine criteriile de evaluare a portofoliului și, în plus, motivația lor pentru realizarea unor produse de calitate va fi mult sporită.

Aprecierea holistică a unui portofoliu reprezintă cea mai bună modalitate de evaluare în cazul acestui instrument. Aceasta se bazează pe impresia generală asupra performanței elevilor sau asupra produselor realizate, luând în considerare elementele individuale componente ale portofoliului, corelată fiind cu rubricile de scoruri pe diferite niveluri și trepte. Scara utilizată poate să cuprindă calificative (de exemplu, „excelent”, „bun”, „acceptabil” și „inacceptabil”), simboluri numerice (de exemplu, se pot folosi cifrele de la 4 la 1 sau de la 5 la 1 etc.) etc.

Exemplul prezentat anterior, aplicabil în cazul limbilor moderne, este un tip de portofoliu care vizează majoritatea capacităților și abilităților implicate în înțelegerea și învățarea unei discipline. Pot fi, însă, concepute și portofolii centrate pe un anumit tip de competență, spre exemplu, citire sau scriere în cadrul disciplinei limbă maternă, și care pot cuprinde sarcini de lucru variate, subsumate aceleiași competențe, pentru a pune în valoare nivelul ei de stăpânire de către elev, progresele de la un moment la altul în evoluția școlară a acestuia.

Portofoliul ca instrument în evaluarea curentă își dovedește utilitatea furnizând informații esențiale profesorului pe baza cărora acesta își poate întemeia o judecată de valoare validă și pertinentă asupra performanței elevului pe o perioadă mai lungă de timp. Fiind un instrument complex și integrator, portofoliul reunește cele mai bune rezultate și produse ale activității elevului, oferind o imagine clară asupra evoluției în timp a acestuia, reflectând motivația pentru învățare și constituind în același timp o modalitate eficientă de comunicare a rezultatelor școlare și a progreselor înregistrate atât pentru elev, cât și pentru părinți sau alte persoane interesate în acest sens.

3.2.5. Autoevaluarea

Prin informațiile pe care le furnizează, autoevaluarea are un rol esențial în întregirea imaginii elevului din perspectiva judecății de valoare pe care o emite profesorul.

Pentru ca evaluarea să fie resimțită de către elev ca având efect formativ, raportându-se la diferite competențe în funcție de progresul realizat și de dificultățile pe care le are de depășit, este foarte utilă formarea și exersarea la elevi a capacității de **autoevaluare**. Elevii au nevoie să știe cât mai multe lucruri despre ei înșiși, despre dimensiunile personalității lor și despre manifestările lor comportamentale. Acest fapt are multiple implicații în plan motivațional și atitudinal.

Ca și profesorul care conduce activitatea, elevul aflat în situația de învățare are nevoie de anumite puncte de referință care să-i definească rolul, sarcina, natura și direcțiile activității sale, ajutându-l să conștientizeze progresele și achizițiile făcute, să-și elaboreze disciplina proprie de lucru, să se poată situa personal în raport cu exigențele de învățare.

Câteva **condiții** necesare pentru formarea capacității de autoevaluare la elevi:

(i) prezentarea încă din debutul activității sau a unei sarcini de lucru a obiectivelor curriculare și de evaluare pe care trebuie să le atingă elevii; încurajarea elevilor pentru a-și pune întrebări legate de modul de rezolvare a unei sarcini de lucru și de efectele formative ale acesteia și pentru a răspunde în scris la acestea;

(ii) încurajarea elevilor pentru a-și pune întrebări legate de modul de rezolvare a unei sarcini de lucru și de efectele formative ale acesteia și pentru a răspunde în scris la acestea;

(iii) stimularea evaluării în cadrul grupului;

(iv) completarea la sfârșitul unei sarcini de lucru importante a unui chestionar care să cuprindă întrebări de tipul celor prezentate în continuare ca exemplu.

Autoevaluarea comportamentelor din sfera domeniului cognitiv, dar și afectiv se poate realiza prin (vezi și Stoica, Musteață, 1997, pp. 1 18-1 19):

- **chestionare** (elevilor li se cere să dea răspunsuri deschise la întrebări);
- **scări de clasificare etc.**

Exemple:

<p>I. Chestionar:</p> <p>1. Care sunt etapele pe care le-ai parcurs în vederea rezolvării efective a sarcinii de lucru?</p> <p>.....</p> <p>.....</p> <p>2. Prin rezolvarea acestei sarcini am învățat:</p> <p>a).....</p> <p>b).....</p> <p>c).....</p> <p>3. Dificultățile pe care le-am întâmpinat au fost următoarele:</p> <p>a).....</p> <p>b).....</p> <p>c).....</p> <p>4. Cred că mi-aș putea îmbunătăți performanța dacă:</p> <p>a).....</p> <p>b).....</p> <p>5. Cred că activitatea mea ar putea fi apreciată ca fiind:</p>
--

.....					
II. Scara de clasificare:					
Fișa de autoevaluare	Slab	Mediu	Bun	F.bun	Excelent
Performanțe școlare	-	-	-	-	-
Motivație și interes constant manifestat prin realizarea diferitelor produse ale activității	-	-	-	-	-
Spirit de inițiativă, independență	-	-	-	-	-
Spirit de cooperare în realizarea unor produse în cadrul activității de grup	-	-	-	-	-
Asumarea responsabilităților în realizarea sarcinilor de lucru curente	-	-	-	-	-
Participarea la discuții în cadrul activității pe grupuri de lucru	-	-	-	-	-
Disciplina de lucru	-	-	-	-	-
Progresul realizat	-	-	-	-	-
Comportamentul general	-	-	-	-	-

Prin utilizarea acestor instrumente în mod constant se poate surprinde în timp evoluția elevilor în planul maturizării psihoafective și a capacității de autoevaluare. Este vorba de o influență intrinsecă generată de *feedback-ul* mult mai intens care se derulează între profesor și elev, în funcție și de specificul disciplinei de studiu.

O altă problemă importantă este aceea a utilității pe care o dăm informațiilor astfel obținute în urma autoevaluării. Pentru a căpăta o semnificație reală, servind formării elevului, ele trebuie integrate și valorificate prin modalități diverse:

- **comparate** cu informațiile obținute de către profesor prin intermediul altor metode complementare;
- **inserate în portofoliul** elevului;
- **prezentate periodic părinților**, alături de alte informații pentru a oferi o

imagine cât mai completă asupra evoluției elevului.

Toate aceste metode complementare de evaluare asigură o alternativă la formulele tradiționale, a căror prezență este preponderentă în activitatea curentă la clasă, oferind alte opțiuni metodologice și instrumentale care îmbogățesc practica evaluativă. Valențele lor formative le recomandă susținut în acest sens.

Este cazul, în special, al investigației, proiectului și portofoliului, care în afara faptului că reprezintă importante instrumente de evaluare, constituie în primul rând sarcini de lucru a căror rezolvare stimulează învățarea de tip euristic.

Dintre **valențele formative** ale metodelor complementare de evaluare pot fi menționate:

- Oportunitatea creată profesorului de a obține noi și importante informații asupra nivelului de pregătire al elevilor săi. Pe baza acestor informații profesorul își fundamentează judecata de valoare, pe care o exprimă într-o apreciere cât mai obiectivă a performanțelor elevilor și a progreselor înregistrate de către aceștia.
- Posibilitatea elevului de a arăta ceea ce știe și, mai ales, ceea ce știe să facă, într-o varietate de contexte și situații.
- Actualizarea permanentă a imaginii asupra performanțelor elevilor, în raport cu abilitățile și capacitățile pe care aceștia le dețin. De asemenea oferă și o imagine cât mai completă asupra profilului general al nivelului de achiziții al elevului, ceea ce este în sprijinul profesorului.

- Asigurarea unui demers interactiv al actului de predare-învățare, adaptat nevoilor de individualizare a sarcinilor de lucru pentru fiecare elev, valorificând și stimulând potențialul creativ și originalitatea acestuia.

- Exersarea abilităților practic-aplicative ale elevilor, asigurând o mai bună clarificare conceptuală și integrare în sistemul noțional a cunoștințelor asimilate, care astfel devin operaționale.

Sumar de idei

Concepte-cheie:

- metodă de evaluare;
- instrument de evaluare;
- metode de evaluare tradiționale;
- metode de evaluare complementare;
- probe orale;
- probe scrise;
- probe practice;
- activități experimentale;
- observarea sistematică a activității și comportamentului elevilor;
- fișa de evaluare;
- scara de clasificare;
- lista de control;
- investigația;
- proiectul;
- portofoliul;
- autoevaluarea.

Idei principale:

- Metoda de evaluare este o cale prin care profesorul oferă elevilor posibilitatea de a demonstra nivelul de stăpânire a cunoștințelor, de formare a diferitelor capacităților testate prin utilizarea unei diversități de instrumente adecvate obiectivului de evaluare propus.

- Relația dintre metoda de evaluare și instrumentul de evaluare este o relație de dependență a instrumentului de metodă, în sensul că instrumentul își subsumează valențele formative și operaționale realizării perspectivei metodologice propuse.

- Metodele de evaluare pot fi împărțite în metode tradiționale și metode complementare.

- Metodele tradiționale de evaluare au căpătat această denumire datorită consacrării lor în timp ca fiind cele mai des utilizate.

- Metodele tradiționale de evaluare trebuie concepute ca realizând un echilibru între probele orale, scrise și cele practice.

- Metodele complementare de evaluare trebuie concepute astfel încât să ofere elevilor suficiente și variate posibilități de a demonstra ceea ce știu (ca ansamblu de cunoștințe), dar, mai ales, ceea ce pot să facă (priceperi, deprinderi, abilități).

- Principalele metode complementare de evaluare, al căror potențial formativ susține individualizarea actului educațional prin sprijinul acordat elevului sunt: observarea sistematică a activității și a comportamentului elevilor; investigația; proiectul; portofoliul; autoevaluarea.

Capitolul IV. Elaborarea și administrarea testelor scrise

În capitolul al III-lea au fost prezentate o serie de metode și instrumente de evaluare în vederea utilizării lor ca mijloace de obținere a informațiilor privind performanțele școlare. Deoarece metoda testelor scrise furnizează un echilibru foarte bun între scopul, calitățile și timpul lor de elaborare, precum și o mai accentuată obiectivitate în notare în comparație cu alte metode, acest capitol va fi consacrat proiectării, administrării, corectării și notării testelor, ca și interpretării rezultatelor obținute în urma aplicării acestor instrumente de evaluare. În plus, extemporalele, lucrările de control semestriale (tezele), testele administrate în cadrul examenelor și evaluărilor naționale rămân cele mai utilizate modalități de apreciere curentă și de examinare. Pentru cititorul mai puțin avizat, elaborarea testelor scrise pare un lucru banal și ușor de realizat. La o privire mai atentă, această concepție nu este conformă cu realitatea, cel puțin datorită câtorva motive pe care le prezentăm în continuare.

În primul rând, pentru a realiza o evaluare relevantă și eficace testele trebuie să evalueze mai puțin cunoștințele acumulate și mai mult aplicarea acestor cunoștințe în situații similare sau în situații noi (rezolvarea de probleme), deoarece numai în acest mod se produce o învățare solidă. O cercetare realizată în 1993 de către Galina Kovalyova arată modul în care înțeleg unele țări să utilizeze instrumentele de evaluare, în general și testele scrise, în particular (vezi schema de la pagina următoare). Se observă că, nu întâmplător, țările dezvoltate economic utilizează testele preponderent în situațiile de tip rezolvare de probleme.

În al doilea rând, orice instrument de evaluare trebuie să îndeplinească anumite exigențe de elaborare, adică anumite „calități tehnice”, în vederea atingerii scopului pentru care acesta a fost proiectat. Principalele calități vor fi descrise și exemplificate în paragraful următor.

În al treilea rând, orice test este compus dintr-un număr de itemi care, pe de o parte au reguli precise de elaborare (vezi capitolul al IV-lea), iar pe de altă parte sunt selectați pe baza unei matrice de specificații. Aceste motive conduc la concluzia că s-a dezvoltat o metodologie distinctă de elaborarea a testelor scrise, pe care o prezentăm pe parcursul acestui capitol. Există două unghiuri de vedere privind utilizarea acestei metodologii. Astfel, în situația proiectării unui test pentru o evaluare externă (examen, evaluare națională pe eșantion reprezentativ de elevi, olimpiadă școlară ș.a.), metodologia trebuie respectată în totalitate. Pe de altă parte, dacă testul este elaborat de către profesori în scopul evaluării curente, atunci numai elementele esențiale ale acesteia trebuie urmărite.

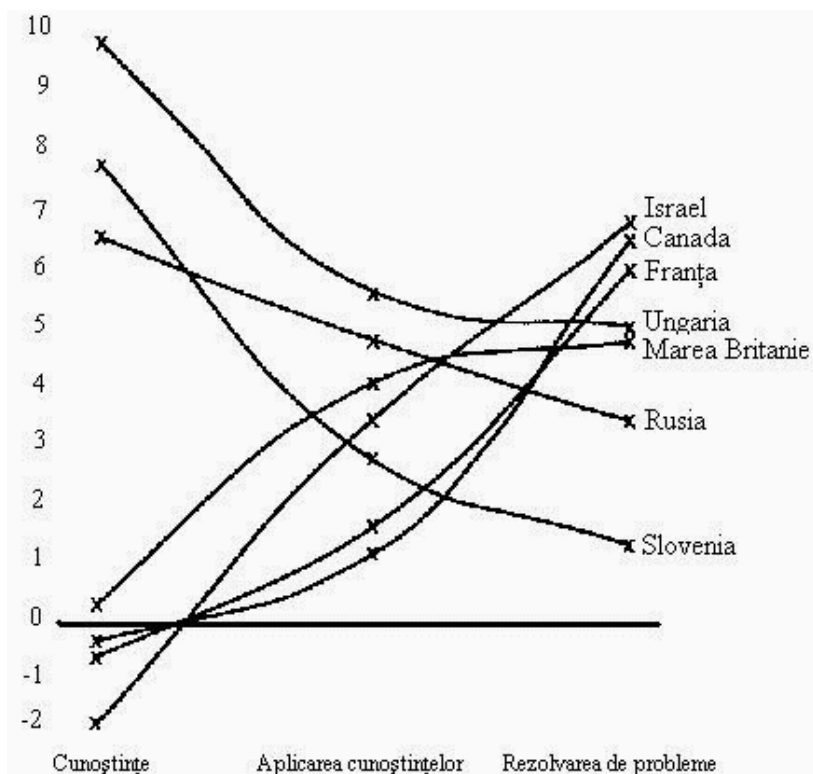


Figura 1. Galina Kovalyova. "Evaluarea comparativă la matematică și științe", Academia rusă de educație, 1993

4.1 Calitățile instrumentelor de evaluare

Evaluarea presupune măsurări, analize și judecăți de valoare. Așa cum un cântar calibrat corespunzător măsoară cu acuratețe greutatea sau un termometru bun indică fidel temperatura, tot așa pentru a efectua măsurători educaționale corecte trebuie să dispunem de instrumente de evaluare care au anumite calități. În acest fel, rezultatele școlare vor fi credibile și vor avea semnificație nu numai pentru evaluatori, dar și pentru evaluați, factori de decizie și societate. Principalele calități ale unui instrument de evaluare — deci implicit ale unui test scris — sunt: **validitatea, fidelitatea, obiectivitatea și aplicabilitatea.**

4.1.1 Validitatea

Conform Ausubel și Robinson, (1981, p.678), validitatea se referă la „faptul dacă testul măsoară ceea ce este destinat să măsoare”. La prima vedere, acest enunț general pare evident, dar semnificațiile sale profunde vor deveni mai clare pe parcursul descrierii tipurilor de validitate. În plus, în cadrul unui test nu trebuie să existe elemente de interferență. De exemplu, un test care evaluează capacitățile matematice nu poate fi scris într-un limbaj greu accesibil vârstei elevilor, deoarece, în acest fel ar test a în primul rând abilitățile de citire și abia după aceea cele matematice. În literatura pedagogică au fost evidențiate următoarele tipuri de validitate:

• **Validitatea de conținut** exprimă măsura în care testul acoperă uniform elementele de conținut majore pe care le testează. Aprecierea validității de conținut se face de către experți, prin estimarea concordanței dintre itemii testului și rezultatele învățării specificate prin obiectivele de evaluare definite. Spre exemplu, dacă am ales un eșantion reprezentativ de obiective de evaluare (80%) pe care dorim să le evaluăm printr-un test, întrebarea la care trebuie să răspundem este: măsoară acesta același eșantion de obiective (și, deci, implicit rezultatele învățării specificate prin aceste obiective)?

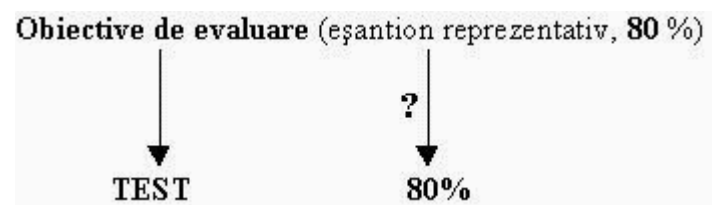


Figura 2. Validitatea de conținut

• **Validitatea de construct** exprimă acuratețea cu care testul măsoară un anumit „construct” (de exemplu: „inteligența”, „succesul școlar”, „motivația”). De pildă, vrem ca un test de inteligență să măsoare inteligența și nu, spre exemplu, creativitatea.

• **Validitatea concurentă** se referă la concordanța dintre rezultatele obținute de elev la un test și „unele comportamente similare”. (Ausubel, D. și Robinson, F., 1981, p.679). De exemplu, dacă un elev care a obținut rezultate bune la un test ce evaluează capacitatea de a utiliza cele patru operații aritmetice nu face greșeli de calcul la un test de tip rezolvare de probleme, în care sunt folosite aceste operații, spunem că primul test are validitate concurentă.

• **Validitatea predictivă** se referă la măsura în care testul face prognoza rezultatelor viitoare ale elevului. Evident, nu putem să evaluăm rezultatele viitoare. De aceea, utilizăm rezultatele prezente pentru a realiza o bună predicție a performanțelor școlare. Pentru estimarea acestui tip de validitate se calculează coeficientul de corelație dintre predictor (exemplu: rezultatele obținute de elevi la examenul de Capacitate la o disciplină) și un criteriu (exemplu: rezultatele acelorași elevi, la aceeași disciplină după primul an de liceu).

• **Validitatea de fațadă („Face Validity”)** exprimă măsura în care testul este relevant și important pentru cei ce sunt testați. De exemplu, un test care evaluează conținuturi de dimensiuni diferite trebuie să exprime acest fapt prin numărul inegal de itemi care testează conținuturile respective.

În tabelul următor sunt prezentate tipurile de validitate necesare diferitelor tipuri de teste:

Validitate Tipuri de teste	concurentă	de construct	de conținut	de fațadă	predictivă
aptitudini		X		X	X
plasament	X		X		X
diagnostic			X		X
de progres	X		X		X

Tabelul 1. Corespondența dintre tipurile de validitate și tipurile de teste

Printre factorii care influențează validitatea unui test menționăm:

- indicațiile neclare;
- nivelul de dificultate necorespunzător al itemilor;
- itemii de calitate slabă;
- lungimea testului (testul prea scurt);
- administrarea și corectarea necorespunzătoare a testului;
- caracteristicile grupului cărui i se administrează testul.

4.1.2 Fidelitatea

Fidelitatea reprezintă calitatea unui test de a produce rezultate constante în cursul aplicării sale repetate. După cum un bun termometru medical arată pentru același om sănătos aceeași temperatură - sau diferențe minime - la diferite momente ale zilei, la fel dorim ca un test aplicat la un anumit interval de timp, în condiții identice, aceluiași grup de elevi, să furnizeze aceleași rezultate (sau rezultate apropiate). Dacă ne referim la un test normativ, acesta nu trebuie să schimbe ordinea candidaților la o aplicare ulterioară. Testele normative au o fidelitate mai ridicată decât testele criteriale, deoarece primele conțin un număr mai mare de itemi și generează scoruri într-un interval mai larg. În același timp, cu cât miza unui test este mai mare, cu atât fidelitatea sa crește. Acest fapt este prezentat în schema următoare:

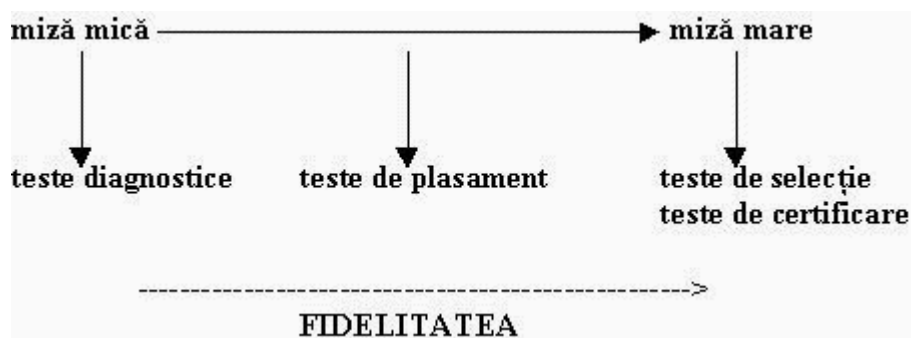


Figura 3. Creșterea fidelității în raport cu miza testului

În acest context, este important ca testele utilizate în examenele naționale să aibă o fidelitate ridicată. Mai mult, această fidelitate ridicată trebuie să se păstreze de la un an la altul pentru ca rezultatele obținute și, în consecință, diplomele eliberate să fie comparabile. Pentru măsurarea fidelității unui test se folosesc metodele descrise succint în tabelul următor (Gronlund N., 1981, p. 96).

Printre factorii care influențează validitatea menționăm:

- Lungimea testului. Cu cât testul este mai lung, cu atât crește fidelitatea sa.
- Dispersia scorurilor. Cu cât împrăștierea scorurilor este mai mare, cu atât testul este mai fidel.
- Obiectivitatea testului. Testul format din itemi obiectivi are o fidelitate mare.
- Schema de notare. O schemă de notare ambiguă scade fidelitatea testului.
- Tipul testului. Cu cât miza testului este mai mare cu atât crește fidelitatea.

Nr.	Metoda	Tipul de fidelitate	Procedeul
1	Test-retest	Stabilitate	Acest test se administrează de două ori aceluiași grup la un interval de timp de câteva minute până la câteva luni.
2	Forme echivalente	Echivalență	La un interval scurt de timp se administrează două forme de test echivalente la același grup de elevi.
3	Test-retest cu forme echivalente	Stabilitate și Echivalență	Se administrează aceluiași grup două forme ale aceluiași test la un interval de timp scurt.
4	Metoda înjumătățirii	Consistența internă	Se aplică testul o singură dată. Se împarte testul în două jumătăți echivalente (de exemplu: itemi pari și impari). Se calculează coeficientul de fidelitate pentru o jumătate de test, cu ajutorul coeficientului de corecție. Se utilizează Spearman-Brown pentru a estima fidelitatea întregului test: Fidelitate test = $\frac{2 \cdot \text{fidelitatea unei jumătăți}}{1 + \text{fidelitatea unei jumătăți}}$
5	Kuder-Richardson	Consistența internă	Se administrează testul o singură dată. Se calculează scorul total și se aplică formula Kuder-Richardson: $KR = \frac{k}{k-1} \left(1 - \frac{nk(k - \bar{x})}{kQ^2} \right)$ unde k = nr. de itemi ai testului \bar{x} : media aritmetică Q = deviația standard

Tabelul 2. Metode de măsurarea fidelității

Metodele care utilizează formele echivalente de test consumă mai multe resurse (pentru proiectarea și administrarea testelor), dar sunt mai relevante în ceea ce privește coeficientul de fidelitate calculat.

Legătura validitate-fidelitate

- Un test poate fi fidel fără a fi valid, deoarece testul poate măsura altceva decât a fost destinat să măsoare.
- Dacă un test nu este fidel, acesta nu este nici valid.
- Fidelitatea este o condiție necesară, dar nu și suficientă pentru validitate.
- Validitatea are o importanță mai mare decât fidelitatea.

Din observațiile de mai sus se desprinde concluzia că în încercarea de a crește fidelitatea unui test trebuie să avem grijă să nu îi scădem validitatea. În acest context, pare, nu de puține ori, că există un „conflict” între validitate și fidelitate în sensul că o creștere a validității conduce la o scădere a fidelității și viceversa. Pentru a rezolva acest conflict trebuie să avem în vedere scopul pentru care am proiectat testul, dar, ca o regulă generală, validitatea trebuie să prevaleze fidelității.

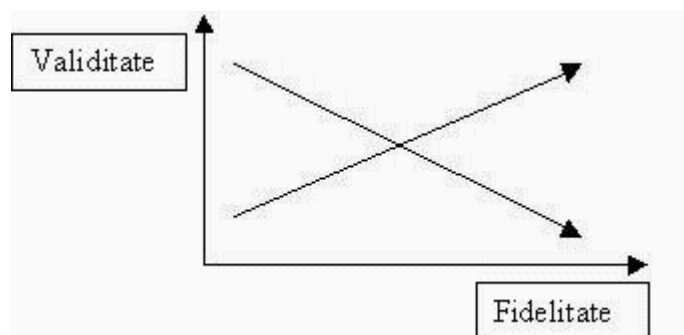


Figura 4. Legătura validitate-fidelitate

4.1.3 Obiectivitatea

Obiectivitatea reprezintă gradul de concordanță între aprecierile făcute de către evaluatori independenți în ceea ce privește un răspuns bun pentru fiecare din itemii testului. Cu alte cuvinte, un test are calitatea de a fi obiectiv, dacă toți cei care îl corectează cad de acord asupra aceluiași răspuns corect pentru fiecare item al testului. Testele cu o foarte bună obiectivitate sunt testele standardizate.

4.1.4 Aplicabilitatea

Aplicabilitatea reprezintă calitatea testului de a fi administrat și interpretat cu ușurință. Criteriile de selectare a testelor cu o bună aplicabilitate sunt (vezi Ausubel și Robinson, 1981, p. 687):

- importanța conținutului pe care testul îl măsoară;
- concordanța dintre forma și conținutul testului și nivelul de vârstă al elevului;
- costul și timpul necesare pentru administrarea testului;
- obiectivitatea în notare și interpretarea rezultatelor.

4.2 Tipuri de teste și scopurile lor

Există mai multe modalități de clasificare a testelor scrise. În continuare vom prezenta două dintre acestea.

4.2.1 Prima modalitate.

În literatura de specialitate sunt enumerate câteva criterii, pe baza cărora sunt clasificate testele: obiectivitatea, utilitatea, specificitatea, aplicabilitatea. Pentru simplitate vom prezenta o serie de tipuri de teste sub forma de perechi formate din elemente opuse.

•Teste de cunoștințe - Teste de aptitudini

Testele de cunoștințe evaluează conținuturi deja parcurse, vizând cunoștințe, priceperi, deprinderi și capacități corespunzătoare acelor conținuturi. Testele de aptitudini au în vedere abilitățile generale ale elevului și nu se referă la un conținut anume.

•Teste criteriale - Teste normative

Testele criteriale presupun aprecierea rezultatelor elevului în raport cu criteriile de performanță anterior stabilite. În plus, testele normative au ca scop și compararea rezultatelor elevului cu cele ale unui grup de referință. Testele normative presupun o ierarhizare a elevilor. Problemele legate de aceste două tipuri de teste le vom *dezvolta într-un paragraf separat*.

•Teste formative - Teste sumative

Scopul testelor formative este de a urmări periodic progresul școlar și, în consecință, de a oferi *feedback-ul* necesar profesorului. Un caz special îl reprezintă testele diagnostice care sunt specifice depistării lacunelor, identificării dificultăților de învățare ale elevului și a modalităților de remediere a acestora. Testele sumative sunt administrate la sfârșitul unei perioade lungi de instruire - semestru, an școlar, ciclu de învățământ - și au ca principal scop notarea elevului. Cele două tipuri de teste furnizează un *feedback* formativ, respectiv sumativ (vezi Capitolul I).

•Teste punctuale - Teste integrative

Testele punctuale conțin itemi care se referă la un aspect izolat al conținutului supus învățării, în timp ce testele integrative sunt formate dintr-un număr mai mic de itemi, dar care — fiecare în parte — evaluează mai multe cunoștințe, priceperi și capacități.

• Teste obiective - Teste subiective

Diferența dintre cele două tipuri de teste constă în gradul de obiectivitate a notării. Cu alte cuvinte, primele sunt teste care conțin itemi ce permit o notare obiectivă, în timp ce testele subiective sunt formate din itemi semiobiectivi și subiectivi, itemi care prin modul de construcție introduc o doză de subiectivitate în corectare și notare.

• Teste inițiale - Teste finale

În primul caz, nivelul performanțelor este evaluat înaintea unui program de instruire, în timp ce în al doilea caz acest nivel este măsurat la încheierea programului de instruire. Un caz special de teste

inițiale îl reprezintă testele de plasament, proiectate cu scopul selectării elevilor pentru un program special de instruire (pentru o limbă străină, o disciplină artistică ș.a.).

Să remarcăm totuși faptul că perechile de teste prezentate mai sus nu sunt total disjuncte. Ele prezintă, așa cum aminteam, elemente contrare, dar au și elemente comune. O altă observație este aceea că un test poate face parte din mai mult de-cât o categorie. De exemplu, un test de cunoștințe poate fi în același timp normativ, integrativ și obiectiv.

4.2.2 A doua modalitate

O altă clasificare a testelor se realizează în funcție de cine le proiectează. Din acest punct de vedere testele se împart în două categorii:

- Testele elaborate de profesori;
- Teste elaborate de instituții specializate (de exemplu teste standardizate).

Testele elaborate de profesori se pot prezenta sub o varietate de forme: de la simple extemporale, la teze și chiar la teste de plasament sau de selecție (pentru olimpiadele școlare). Principala lor fragilitate constă în calitățile tehnice scăzute, în special a unei fidelități necorespunzătoare. Așa cum aminteam la începutul capitolului, nimeni nu are pretenția ca un profesor — care are și alte numeroase sarcini didactice — să elaboreze **frecvent** teste de bună calitate, ci să înțeleagă și să respecte cerințele minimale de proiectare ale acestora. În plus, fiecare profesor ar trebui ca la anumite intervale de timp să încerce:

- să elaboreze itemii testelor respectând regulile prezentate în capitolul următor;
- să asambleze acești itemi în teste pe baza unei matrice de specificații;
- să construiască o schemă de corectare care să permită o notare obiectivă a testelor.

Acest exercițiu este cu atât mai important și util, cu cât marea majoritate a profesorilor participă ca evaluatori în examenele naționale.

Mulți autori consideră că **testele standardizate** evaluează, în special, procesele intelectuale inferioare și numai ocazional pe cele de nivel superior. Un studiu realizat de Romberg în SUA (Stenmark, I., 1991) pe cele mai cunoscute șase teste standardizate de matematică a arătat că, la nivelul clasei a VIII-a, doar 1 % dintre itemi erau de tip rezolvare de probleme, în timp ce 77% se refereau la simple calcule.

Testele standardizate sunt alcătuite din itemi obiectivi și, mai rar, semiobiectivi, ce sunt selectați dintr-o bancă de itemi în care sunt grupați în funcție de disciplină, obiective, nivel de dificultate ș.a. (vezi [4.5](#)). În vederea elaborării testelor standardizate este nevoie să fie creați un număr mare de itemi paraleli, care să măsoare același obiectiv, fapt ce impune participarea mai multor specialiști. Acestora li se furnizează informații clare privind caracteristicile itemilor ce trebuie elaborați în vederea asigurării comparabilității itemilor de același tip. Mai mult, administrarea testelor standardizate este foarte riguroasă, aceasta desfășurându-se în aceleași condiții (standard) pentru toți candidații.

Unul dintre cele mai cunoscute teste standardizate este **GRE** (*Graduate Record Examination*) care este utilizat în SUA pentru admiterea studenților în programe de studiu avansate (master sau doctorat). **GRE**-ul este, de fapt, o baterie de teste ce evaluează concepte și capacități care au fost achiziționate și formate de-a lungul unei perioade mai mari de timp. Testul nu este centrat pe o singură disciplină — ceea ce îl deosebește fundamental de testele de cunoștințe — și, în consecință, constituie criteriul obligatoriu de admitere într-un program de studiu, indiferent de specializare. **GRE**-ul este format din trei secțiuni distincte: o parte „verbală”, o parte „cantitativă” și una „analitică”. Secțiunea verbală — elaborată de specialiști în științe umaniste — verifică abilitățile de limba engleză (gramatică, vocabular, înțelegerea și interpretarea conținuturilor unor texte etc.), secțiunea cantitativă — realizată de experți în matematică — testează capacitățile matematice (la aritmetică, algebră, geometrie, trigonometrie și statistică), iar secțiunea analitică — proiectată de specialiști în lingvistică și filosofie — conține itemi ce evaluează gândirea logică a studenților, capacitatea de analiză și sinteză, creativitatea etc.

Am subliniat faptul că mulți autori cred că testele standardizate evaluează într-o măsură mică nivelurile cognitive superioare. **GRE**-ul, ca și celelalte teste de aptitudini, face excepție.

Testele standardizate se clasifică în **teste de cunoștințe (criteriale și normative)** și **teste de aptitudini (normative)**. Vom descrie în continuare aceste tipuri de teste.

Principala diferență între testele **criteriale** și cele **normative** constă în modul în care sunt interpretate rezultatele candidaților. Primele comunică performanța absolută a unui candidat (în raport cu obiectivele stabilite), în timp ce testele normative comunică performanța relativă a fiecărui candidat, adică performanța sa în raport cu ceilalți candidați. În tabelul următor sunt prezentate și celelalte deosebiri:

Teste criteriale	Teste normative
Sunt utilizate pentru certificare.	Sunt utilizate pentru selecție.
Evaluează un număr mai mic de obiective și conținuturi, dar în profunzime mai mare.	Evaluează un număr mai mare de obiective și conținuturi, asigurând o bună acoperire a curriculum-ului.
Sunt utilizate și în scop diagnostic.	
	Toți itemii sunt de dificultate medie, în vederea unei bune discriminări a rezultatelor candidaților.

Tabelul 3. O comparație între testele criteriale și testele normative

Testele de aptitudini au ca principal scop prognoza rezultatelor viitoare ale elevilor. Aceste teste nu evaluează un anumit conținut curricular, ci diferite aptitudini pe care elevul le posedă la un moment dat și care pot favoriza, în viitor, performanțele sale școlare. Teste de aptitudini sunt, în general, identificate cu testele de inteligență. Confuzia provine din două motive:

- mulți cred că inteligența este ceva moștenit (înnăscut);
- există numeroase controverse privind înțelesul noțiunii de inteligență și a elementelor ce pot fi incluse în acest concept.

În concluzie, testele standardizate sunt distincte față de testele elaborate de profesori. Printre aceste deosebiri enumerăm (Gronlund, 1981, p.306):

- Itemii testelor au calități tehnice superioare. Aceștia au fost proiectați pe baza matricelor de specificații, pretestați și selectați pe baza analizei de item.
- Indicațiile privind administrarea și corectarea testelor sunt atât de precise, încât procedurile sunt aceleași (standard) pentru diferiți utilizatori.
- Norme — bazate pe grupuri reprezentative de elevi — sunt furnizate pentru a ajuta în interpretarea scorurilor. Normele sunt prezentate pe vârste, la nivel național sau regional.
- Forme echivalente de teste sunt elaborate.
- Un manual al testului și alte materiale ajutătoare sunt elaborate pentru a servi drept ghid în administrarea și corectarea testului, pentru evaluarea calităților sale, pentru interpretarea și utilizarea rezultatelor.

Un alt tip important de teste sunt **testele de progres școlar**. Acestea evaluează stadiul la care elevul a ajuns în realizarea obiectivelor instruirii, deci au un scop diagnostic. Prezentăm în continuare câteva teste de progres școlar pentru domenii specifice (vezi Gronlund, 1981, p. 313 și următoarele):

• Teste orientate spre conținut

Testele sunt destinate să măsoare conținutul unei anumite discipline: matematică, științe, limbi străine, științe sociale ș.a.

• Teste de înțelegere a mesajului scris

Testele evaluează, de obicei, vocabularul, înțelegerea unui text și viteza de citire. De exemplu:

- identificarea înțelesului unor cuvinte;
- identificarea înțelesului cuvintelor utilizate într-un anumit context;
- identificarea detaliilor care se găsesc într-un text;
- identificarea ideilor conținute într-un text.

• Teste diagnostice

Testele furnizează detalii privind dificultățile de învățare ale elevilor. Principalele lor caracteristici sunt:

- testează numai acele obiective importante pentru diagnoză;
- conțin un număr mare de itemi, care măsoară același obiectiv;
- dificultatea testului este scăzută, pentru a face o bună discriminare între elevii cu dificultăți de învățare.

Cele mai multe teste diagnostice sunt elaborate pentru matematică și citire.

Exemplu: Testul diagnostic de matematică Stanford conține următoarele părți (în paranteze s-a trecut numărul de itemi):

Testul 1: Numere

- Numere naturale (21 itemi)
- Numere raționale (3 itemi)
- Operații și proprietăți (12 itemi)

Testul 2: Calcule

- Adunarea numerelor întregi (6 itemi)
- Scăderea numerelor întregi (12 itemi)
- Înmulțirea numerelor întregi (12 itemi)
- Împărțirea numerelor întregi (9 itemi)

Testul 3: Aplicații

- Rezolvarea de probleme (12 itemi)
- Citirea și interpretarea tabelelor și graficelor (9 itemi)
- Geometrie și măsurarea (9 itemi)

Pentru niveluri de vârstă mai mari, testul 2 conține și: fracții, numere zecimale și procente.

Am descris mai sus tipurile de teste și scopurile lor. Să vedem acum în ce relație se află principalele scopuri cu diferite strategii de evaluare a performanțelor școlare (adaptare după Somerset, NCA, 1993).

Scop	Evaluare curentă	Examene	Evaluare națională (pe eșantion reprezentativ)
Evaluarea performanțelor pe parcursul școlarizării.	DA	NU	DA
Monitorizarea în timp a performanțelor la nivel național .	NU	Nu, deoarece se utilizează alte teste de la un examen la altul și de la un an la altul.	Da, cu condiția să fie asigurat secretul testelor utilizate la diferite intervale de timp.
Furnizarea unui <i>feedback</i> diagnostic.	DA	Da, cu toate că este foarte dificil ca anumite conținuturi să poată fi evaluate în profunzime.	Da, cu condiția ca testele să nu fie păstrate secret după administrare.
Influențarea procesului de predare-învățare prin efectul <i>backwash</i> .	NU	Da, mai ales dacă miza examenului este mare.	Nu, dar rezultatele pot oferi sprijin pentru îmbunătățirea procesului de predare-învățare.

Tabelul 4. *Relația dintre scopurile testelor și strategiile de evaluare*

În proiectarea unui test trebuie avute în vedere, printre altele, următoarele etape:

- Determinarea tipului de test.
- Proiectarea matricei de specificații.

- (iii) Definirea obiectivelor de evaluare.
- (iv) Construirea itemilor.
- (v) Elaborarea schemei de notare.
- (vi) Pilotarea și revizuirea testelor, precum și a schemei de notare.
- (vii) Administrarea testelor.
- (viii) Corectarea și analiza rezultatelor.

În termenii de produse și activități, elaborarea testelor se realizează schematic în modul următor:

Produse	Activități
<ul style="list-style-type: none"> matrice de specificații 	<ul style="list-style-type: none"> Activitate (de tip expert) în grupurile de lucru pe discipline
↓	↓
<ul style="list-style-type: none"> obiective de evaluare 	<ul style="list-style-type: none"> Activitate (de tip expert) în grupurile de lucru pe discipline;
↓	↓
<ul style="list-style-type: none"> itemi 	<ul style="list-style-type: none"> Activitate (de tip expert) în grupurile de lucru pe discipline; pretestarea itemilor;
↓	↓
<ul style="list-style-type: none"> teste și scheme de lucru 	<ul style="list-style-type: none"> Activitate instituțională: corectarea testelor, analiza și interpretarea rezultatelor
↓	
<ul style="list-style-type: none"> rapoarte pentru diferite categorii de beneficiari 	

Tabelul 5. *Elaborarea testelor (activități și produse)*

4.3 Matricea de specificații

Odată tipul de test determinat — de exemplu, normativ sau criterial — avem nevoie de un procedeu prin care să fim siguri că testul măsoară obiectivele educaționale definite anterior și are o bună validitate de conținut. Unul dintre cele mai utilizate procedee care să servească acest scop este construirea **matricei de specificații** pentru testul respectiv. Pe liniile matricei sunt enunțate conținuturile testate, iar coloanele conțin nivelurile cognitive la care dorim să măsurăm aceste conținuturi (de exemplu: cunoaștere, înțelegere, aplicare, analiză și sinteză, conform taxonomiei obiectivelor educaționale a lui Bloom, B.S.). În unele lucrări de specialitate taxonomia lui Bloom este modificată pentru a o „adapta caracteristicilor specifice performanței intelectuale din fiecare disciplină în parte (Ausubel, și Robinson, 1981, p.60). Nu este un lucru neobișnuit nici să se combine categorii din diferite taxonomii. Este chiar indicat să se întâmple acest fapt deoarece toate taxonomiile (inclusiv cea a lui B. S. Bloom) au limite și nu pot satisface exigențele evaluării la orice disciplină (Stoica, A., 2000, p. 70)”. Exemplu:

	Cunoaștere și înțelegere	Aplicarea	Rezolvare de probleme	Total (%)
C ₁	6	9	15	30
C ₂	8	12	20	40
C ₃	4	6	10	20
C ₄	2	3	5	10
Total (%)	20	30	50	100

Să presupunem că dorim să testăm trei tipuri de comportamente/domenii-Cunoaștere și înțelegere, Aplicarea, Rezolvare de probleme-la 4 elemente de conținut-C₁, C₂, C₃, C₄. Vom determina apoi ponderea pe care fiecare domeniu și element de conținut o va avea în cadrul testului. Astfel, putem atribui celor trei domenii ponderile 20%, 30% și 50%, respectiv (ultima linie a matricei), iar conținuturilor C₁, C₂, C₃, C₄ ponderile 30%, 40%, 20% și 10%, respectiv în economia testului proiectat.

Având aceste date, va fi ușor să completăm celulele matricei prin înmulțirea liniilor cu coloanele corespunzătoare. De exemplu, ponderea „Aplicării” în conținutul 1 (C_1) este:

$$30\% \times 30\% = 9\%.$$

De multe ori se obișnuiește ca în locul ponderii să fie trecut direct numărul de itemi din fiecare celulă. Stabilindu-se numărul total de itemi ai testului - de exemplu 40 -, transformarea este simplă. Astfel, numărul itemilor care testează rezolvarea de probleme pentru conținutul 2 va fi:

$$20 / 100 \times 40 = 8 \text{ itemi.}$$

În funcție de scopul și tipul testului pot fi construite matrice de specificație **generale** sau **particulare (detaliat)**.

(i) Matricele cu caracter general însoțesc testele sumative care evaluează elemente mai mari de conținut și domenii cu un grad general de specificitate.

Exemplu (matrice de specificații la Algebră, clasa a VII-a):

Conținut	Nivel			
	Înțelegere	Aplicare	Rezolvare de probleme	Nr. itemi
Funcții				
1. Noțiunea de funcție	X	X		
1.1. Exemple	X	X		
1.2. Reprezentare grafică				
2. Funcții liniare	X	X	X	
2.1. Definiție	X	X		
2.2. Reprezentare grafică				
3. Funcții pătratice		X	X	
4. Alte funcții		X	X	
Nr. itemi				

Pentru fiecare dintre celule matricei marcate prin „x” se vor elabora unul sau mai mulți itemi. Exemplul arată, de asemenea, că pot exista și celule goale.

(ii) O matrice detaliată se obține prin următoarele căi, considerate fie separat, fie împreună:

- împărțirea conținuturilor în subconținuturi;
- împărțirea domeniilor în subdomenii.

În exemplul următor vom utiliza cea de-a doua modalitate:

Domenii	Cunoaștere		Înțelegere	Interpretare	Total
Conținuturi	Simboluri și termeni	Fapte specifice	Influența fiecărui factor asupra caracteristicilor vremii	Hărți meteorologice	
Presiunea atmosferică	2	3	3	3	11
Vântul	4	2	8	2	16
Temperatura	2	2	2	2	8
Umiditatea și precipitațiile	2	1	2	5	10
Norii	2	2	1		5
<i>Total</i>	12	10	16	12	50

Un caz special de detaliere a unei matrice de specificații se referă la includerea în matrice a domeniilor foarte importante pentru conținuturile evaluate, dar care nu pot fi testate în scris (Gronlund, N., 1981, p. 130). Este evident faptul că un singur instrument de evaluare nu poate măsura totul. Prin includerea tuturor obiectivelor pe care dorim să le evaluăm va fi mai clar ce și cât anume poate măsura testul pe care îl proiectăm. În exemplul următor sunt incluse două capacități, care nu pot fi evaluate printr-un test scris. Ponderea acordarea acestora este 25% pentru fiecare.

	Cunoaștere		Înțelegere	Interpretare	Capacitatea de a		Total
	Simboluri și termeni	Fapte specifice	Influența fiecărui factor asupra caracteristicilor vremii	Hărți meteorologice	Utiliza instrumente de măsură	Construi hărți meteorologice	
Presiunea atmosferică	2	3	3	3	Metodă:	Metodă:	11
Vântul	4	2	8	2	Observarea elevilor în timpul utilizării instrumentelor (scară de notare)	Evaluarea hărților construite de elevi (fișă de evaluare)	16
Temperatura	2	2	2	2			8
Umiditatea și precipitațiile	2	1	2	5			10
Norii	2	2	1				5
<i>Total itemi</i>	12	10	16	12			50
<i>Pondere</i>	12%	10%	16%	12%	25%	25%	100%

4.4 Schema de notare

Cu ajutorul matricei de specificații se vor scrie obiectivele de evaluare și apoi itemii relevanți pentru fiecare obiectiv. Pentru a încheia procesul de proiectare a testului, trebuie definită schema de notare, pe baza căreia vor fi corectate și punctate lucrările elevilor. La începutul capitolului am arătat că una dintre modalitățile de a crește fidelitatea unui test este elaborarea unei scheme de notare adecvată. Acest lucru nu este ușor de realizat, necesitând un efort susținut și o experiență bogată. Dificultatea alcătuirii schemei de notare este în raport direct cu tipul de itemi utilizat. Deoarece pentru itemii obiectivi și semiobiectivi notarea nu prezintă dificultăți majore, ne vom îndrepta atenția în cele ce urmează asupra schemei de notare pentru itemi care solicită un răspuns deschis.

Există două modalități principale de proiectare a schemei de notare, indiferent dacă aceasta însoțește un test elaborat de profesor sau un test de examen. Aceste două modalități au în vedere **notarea analitică** și **notarea holistică (globală)**. **Notarea analitică** se aplică, în special, testelor de tip formativ, când principalul scop este acela de a identifica și analiza erorile elevilor cu scopul ameliorării programului de instruire și învățare. Procedeu presupune determinarea principalelor performanțe (unități de răspuns) pe care elevul trebuie să le evidențieze în răspunsul său la fiecare item. Unităților de răspuns li se acordă puncte care, însumate, determină nota (scorul) pentru un anumit item.

Notarea analitică prezintă câteva dezavantaje. În primul rând, oricât de bine am determina lista de performanțe, deoarece nu există „răspunsuri fixe”, ne putem aștepta oricând din partea elevilor la un răspuns corect *neanticipat*. Acest fapt poate crea probleme serioase în cadrul unui examen național, atunci când mii de examinatori utilizează aceeași schemă de notare. În unele țări se încearcă evitarea acestui dezavantaj prin alcătuirea unei scheme de notare provizorii odată cu proiectarea testului și definitivarea acesteia după ce un număr semnificativ de lucrări au fost corectate (aproximativ 20% din numărul total). În al doilea rând, profesorii / examinatorii au nevoie de un efort de concentrare și un timp îndelungat pentru corectarea lucrărilor pe baza schemei de notare. În al treilea rând, pentru a evita erorile de aplicare a schemei de notare este necesar ca cel puțin doi examinatori să corecteze aceeași lucrare.

Notarea holistică (globală) este utilizată atunci când nu este necesar *un feedback* asupra naturii erorilor. Ea constă în formarea de către profesor/examinator a unei impresii despre un răspuns în

totalitatea sa și încadrarea acestui răspuns într-o categorie prestabilită. L. Carey (1988, p.190) enumeră șapte etape în proiectarea unei scheme de notare de tip holistic:

a) Stabilirea categoriilor care vor fi utilizate. De exemplu, răspunsurile pot fi plasate în una din următoarele categorii:

(i) admis, respins;

(ii) bun, satisfăcător, slab;

(iii) 0-2 puncte, 2-4 puncte, 4-6 puncte, 6-8 puncte, 8-10 puncte.

Numărul categoriilor depinde de scopul testului, dar acest număr nu trebuie să fie nici prea mic, nici prea mare. În caz contrar devine dificilă încadrarea răspunsurilor în diferitele categorii predefinite (în general se consideră că cinci este un număr corespunzător).

b) Determinarea criteriilor de evaluare pentru fiecare categorie. De exemplu, ce capacități trebuie să demonstreze elevul pentru ca răspunsul său să fie clasificat drept „excellent”?

c) Citirea rapidă de către examinator a tuturor răspunsurilor testelor și formarea impresiei generale.

d) Încadrarea răspunsurilor în categoriile stabilite.

e) Recitirea testelor plasate în cadrul aceleiași categorii, pentru a face comparații.

f) Reîncadrarea anumitor teste în categorii superioare sau inferioare categoriei în care au fost plasate inițial.

g) Acordarea aceleiași note tuturor testelor încadrate în aceeași categorie. De exemplu, toate lucrările din categoria „bine” vor primi 4 puncte, cele din categoria „foarte bine” - 5 puncte ș.a.m.d.

Nici una din etapele prezentate - cu excepția ultimei - nu sunt ușor de realizat. Pentru c) și d) Thyne, I. M. (1978, p.232-233) propune următoarea abordare: să presupunem că am stabilit cinci categorii, denumite A, B, C, D și E. Profesorul/examinatorul citește rapid, „în linii mari”, fiecare test - dacă numărul acestora este prea mare, se va alege un eșantion de 30-40 teste - pentru a-și forma o impresie generală asupra performanțelor elevilor. După aceea, alege cel mai bun și cel mai slab test și le încadrează în categoriile A, respectiv E. Pasul următor constă în a găsi o lucrare de categoria C, adică la jumătatea „distanței” dintre A și E. În final, alege încă două teze care se situează calitativ între A și C, respectiv între C și E. Aceste cinci teste vor servi drept criterii pentru aprecierea celorlalte. Vom prezenta în continuare un exemplu corespunzător etapelor a) și b), respectiv determinarea categoriilor și a criteriilor de apreciere corespunzătoare fiecărei categorii:

Claritatea argumentării și prezentării		
Categorie	Criterii	Puncte
A	O idee clară străbate întregul răspuns. Un limbaj gramatical și tematic corespunzător a fost utilizat pentru prezentarea argumentelor. În susținerea argumentelor au fost foarte bine utilizate simbolurile, graficele și figurile matematice. Elevul a demonstrat capacitatea de a raționa și a justifica rezultatele obținute.	10
B	Prezentarea este clară, conținând un scris fluent și metode grafice. Părți de demonstrații separate au fost puse împreună. Cititorul este satisfăcut în final, dar fără să perceapă o schimbare/provocare intelectuală.	9
C	Prezentarea este clară, dar un domeniu restrâns de tehnici a fost utilizat. Acestea au fost folosite corect. Materialul este prezentat într-o succesiune logică. Unele elemente au fost explicate, dar altele au fost lăsate fără explicație, pe seama cititorului.	7-8
D	Prezentarea este adecvată și poate fi parcursă cu un oarecare efort. Limbajul, simbolurile, regulile matematice au fost utilizate corect. Există o încercare de organizare a materialului.	5-6
E	Câteva aspecte ale problemei au fost abordate, dar nici unul nu a fost finalizat.	3-4
F	O prezentare și organizare întâmplătoare a activității, foarte dificil de înțeles și urmărit. O serie de elemente separate, fără o legătură nici între ele.	1-2

Tabelul 6. Notare holistică

Principala critică adusă metodei de notare holistică este gradul uneori prea general al criteriilor de evaluare (Stoica, 2000). Din acest motiv devine foarte important punerea în practică a unui sistem de **moderare internă și externă**. Scopul acestui sistem este asigurarea unei fidelități ridicate a testelor. **Moderarea internă** se referă la acordul realizat între profesorii de aceeași disciplină dintr-o școală de a interpreta criteriile în același mod și deci de a asigura pe această cale obiectivitatea și comparabilitatea în notare. Odată stabilit același mod de notare în cadrul școlii se trece la moderarea externă care, pe scurt, înseamnă ca același mod de notare să existe și între școli. Pentru acest lucru este nevoie de evaluatori externi, cum ar fi inspectorii școlari.

Un caz special îl reprezintă sistemul de evaluare, prin calificative, (excelent, foarte bine, bine, suficient, insuficient), introdus în învățământul primar în anul 1998. Noul sistem are la bază criterii unitare pentru evaluarea performanțelor elevilor la nivel național, numite **descriptori de performanță**. În capitolul al VI-lea vom detalia unele aspecte ale reformei evaluării în ciclul primar.

4.5 Bănci de itemi

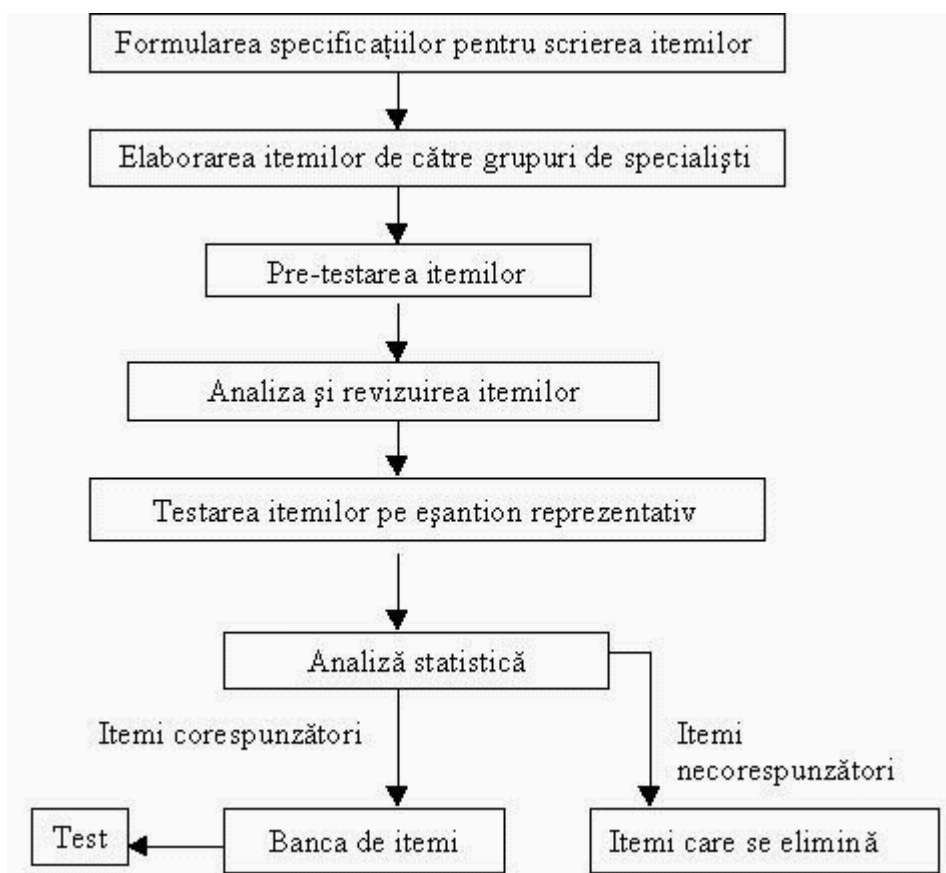
„O bancă de itemi este o colecție de itemi organizați, clasificați și catalogați ca și cărțile în bibliotecă, din care aceștia pot fi aleși cu scopul de a crea teste (Choppin, B., 1987, p. 274). Un avantaj important al băncilor de itemi este acela că sunt utilizate în toate formele de evaluare: formativă, sumativă, internă, externă, inițială, finală ș.a. Importanța băncilor de itemi rezultă din scopurile în care acestea sunt create (vezi și Wood, Stumik, 1969, p. 8):

- (i) existența unor colecții de itemi de bună calitate, pentru diferite discipline, care să fie utilizate atât în evaluările interne, cât și în cele externe;
- (ii) construirea de forme paralele de teste care pot fi aplicate pe grupuri de elevi la momente diferite, cu asigurarea compatibilității rezultatelor;
- (iii) elaborarea, cu mai mare ușurință, a testelor standardizate;
- (iv) familiarizarea cât mai multor profesori cu modul de proiectare a itemilor, în scopul îmbunătățirii acțiunilor evaluative la clasă;
- (v) identificarea celor ce pot scrie itemi de bună calitate.

Itemii sunt clasificați în bănci în funcție de următoarele criterii:

- Clasă (nivel de vârstă)
- Disciplină
- Tip
- Obiectiv evaluat
- Date statistice (de exemplu, indicele de dificultate și indicele de discriminare).

Schema următoare prezintă o modalitate de creare a unei bănci de itemi:



Odată o bancă de itemi construită, testele (chiar forme paralele) pot fi elaborate cu ușurință. Pe măsură ce aceste teste sunt administrate banca trebuie „alimentată” cu noi itemi, urmând procedura descrisă mai sus.

4.6 Proiectarea testelor pentru evaluări externe

În vederea elaborării unor teste pentru evaluări externe (examene, testări naționale ș.a.), care să satisfacă acele calități descrise în paragraful 4.1 trebuie respectată o metodologie ale cărei etape sunt descrise în continuare.

- Obținerea unor informații precise privind scopul testului și resursele pe care elaborarea acestuia le presupune (umane, financiare, materiale și de timp).
- Definirea matricei de specificații.
- Definirea obiectivelor de evaluare ale testului (obiectivele operaționale).
- Scrierea itemilor.
- Pre-experimentarea itemilor, analiza și revizuirea acestora.
- Experimentarea itemilor (pe eșantion reprezentativ), analiza statistică și selectarea itemilor corespunzători din punct de vedere psihometric.
- Asamblarea itemilor în test(e) pe baza matricelor de specificații.

Deoarece în cadrul acestui capitol și al capitolului următor sunt descrise celelalte etape, ne oprim acum asupra experimentării itemilor (vezi și Stoica, A., 2000, p. 72-73). Acest proces trebuie condus cu profesionalism de către instituția organizatoare (Serviciul Național de Evaluare și Examinare, în cazul țării noastre), având în vedere că există mulți participanți la scrierea itemilor, fiecare cu grade diferite de expertiză și experiență în această activitate. Procesul de experimentare conține două etape:

- **Pre-experimentarea** pe un eșantion redus - de obicei câteva clase de elevi - pentru a verifica gradul de adecvare a itemilor la nivelul de vârstă și cerințele curriculum-ului, claritatea, erorile științifice și funcționarea schemei de notare. În această etapă o bună parte dintre itemi necesită o revizuire, iar unii dintre aceștia vor fi chiar eliminați. Pre-experimentarea nu oferă garanția unor itemi de bună calitate, dar este importantă pentru înlăturarea unor erori grave ce pot apărea în elaborarea lor.

- **Experimentarea** propriu-zisă a itemilor pe grupuri mari de elevi, adică pe eșantioane reprezentative ale populației pentru care itemii au fost elaborați. Scopul experimentării este acela de a selecta doar itemii cu calități tehnice superioare, fapt ce poate fi constatat, în special, din analiza statistică a rezultatelor obținute. Ne referim aici la analiza de item - indicii de dificultate, de discriminare, de sensibilitate și de corelație - iar în cazul itemilor obiectivi și la analiza distractorilor. Completată cu judecata de tip expert asupra validității itemilor - concordanța itemi-obiective-conținuturi - etapa experimentării va conduce la reținerea itemilor de foarte bună calitate. Experimentarea este un proces complex care pune cel puțin două probleme majore:

(i) aplicarea testelor pe eșantioane reprezentative de elevi, care presupune probleme de administrare (vezi paragraful următor) și financiare;

(ii) testarea unui număr mare de itemi în raport cu durata limitată de administrare a unui test (2-3 ore) și cu numărul mic de experimentări impus de complexitatea acestui proces. O soluție ar fi gruparea itemilor de același tip și care evaluează același obiectiv și crearea unor grupuri de teste paralele. Va fi suficient ca un singur test din fiecare grup să fie aplicat.

Fie și numai din perspectiva acestor două probleme majore se observă că experimentarea itemilor și, implicit, elaborarea băncilor de itemi — este un proces complex, care se desfășoară pe perioade lungi. Uneori, crearea unei bănci de itemi la o disciplină poate dura ani de zile. Dar odată create, băncile de itemi aduc beneficii considerabile acțiunilor evaluative, atât externe cât și interne.

În vederea selectării itemilor corespunzători pentru un test, ne vom referi, pe scurt, la *analiza statistică de item*, care se va realiza după faza de experimentare. Prin analiza de item se înțelege calculul unor indicatori - pe baza rezultatelor obținute de elevi la un item - și interpretarea lor statistică. Principalii indicatori statistici care se calculează pentru itemii obiectivi - itemi din care sunt în general formate testele normative - sunt indicele de dificultate și indicele de discriminare.

- Indicele de dificultate reprezintă proporția celor care au răspuns corect un item. Deci indicele de dificultate se calculează prin împărțirea numărului de elevi care au rezolvat corect itemul la numărul total de elevi și are valori cuprinse între 0 și 1. Valori mai mici decât 0,20 (item foarte dificil) și mai mari ca 0,80 (item foarte ușor) nu sunt indicate.

- Indicele de discriminare reprezintă corelația dintre rezultatele la un item și un criteriu intern sau extern, de exemplu performanța (scorul) obținută (obținut) de elevi la testul din care face parte itemul respectiv. Cu alte cuvinte, un indice de discriminare bun pentru un item va arăta că elevii care au obținut scoruri mari la test au obținut scoruri mari și la itemul respectiv, iar cei ce au avut scoruri mici la test au obținut scoruri mici inclusiv la acel item. Există mai multe metode de calcul al indicelui de discriminare (D). Una dintre acestea urmează următorii pași:

(i) se scriu în ordine crescătoare rezultatele tuturor elevilor care au susținut testul;

(ii) se calculează indicele de dificultate pentru primii 27% și ultimii 27% dintre elevi;

(iii) valoarea indicelui de discriminare va rezulta din diferența celor doi indici de dificultate.

D ia valori între -1 și 1. Bazându-se pe experimente, C. Croker și J. Algina (1986, p. 315) stabilesc următoarele interpretări ale indicelui de discriminare:

a) dacă $D > 0,40$, itemul este bun;

b) dacă $0,30 < D < 0,40$, este necesară o mică revizuire a itemului;

c) dacă $0,20 < D < 0,30$, este necesară o revizuire foarte atentă a itemului;

d) dacă $D < 0,20$, itemul va fi eliminat.

Prin alegerea unor indici de dificultate și discriminare corespunzători pentru itemii unui test vom asigura, în bună parte, *fidelitatea* testului respectiv.

În cazul *testelor criteriale* problema analizei de item se pune în mod diferit. Indicele de dificultate nu are nici o relevanță deoarece dificultatea este determinată de criteriu/obiectivul pe care îl măsoară.

De exemplu, în situația în care criteriul de performanță specificat de obiectivul de evaluare pe care elevul trebuie să-l realizeze, este unul minimal, itemul va fi și el proiectat la un nivel scăzut de dificultate. Deci, nu este vorba de un alt mod de calcul al indicelui de dificultate (față de testele normative), ci de slaba sa relevanță. Nici indicele de discriminare nu are o importanță prea mare, deoarece la testele criteriale nu urmărim o dispersie a rezultatelor. Mai mult, putem întâlni situația în care toți elevii au rezolvat corect un item, ceea ce este un lucru ideal pentru astfel de teste. Aceasta înseamnă ca indicele de discriminare este 0. Deci, conform criteriilor valorice prezentate mai sus, acest item trebuie eliminat dintr-un test normativ, dar el va fi menținut într-un test criterial (Gronlund, N., 1981, p. 265). Cox și Vargas (1986) consideră că discriminarea trebuie făcută între elevii care au urmat curriculum-ul evaluat și cei care nu l-au urmat încă. Ei propun un nou indicator statistic numit *indice de senzitivitate*, având în acest caz aceeași funcție ca indicele de discriminare pentru testele normative. Indicele de senzitivitate se calculează după cum urmează:

- (i) se aleg aleatoriu două grupuri de elevi: experimental și de control;
- (ii) grupului experimental i se predă conținutul la care se referă testul;
- (iii) se aplică testul ambelor grupuri și se calculează indicii de dificultate;
- (iv) diferența acestor doi indici va reprezenta indicele de senzitivitate.

Valoarea acestuia este cuprinsă între -1 și 1 și este de așteptat să fie cât mai apropiată de 1. Pe baza indicilor statistici și a verificării validității, procesul de selectare a itemilor poate fi considerat încheiat.

4.7 Administrarea testelor scrise

Administrarea testelor reprezintă o etapă vitală în desfășurarea oricărei evaluări externe. Întregul efort de elaborare a testelor și schemelor de notare poate fi compromis printr-o administrare necorespunzătoare. Pe de altă parte, posibilitățile de administrare cât mai corectă (standardizată) a testelor influențează mult modul de proiectare al acestor. Putem vorbi și aici despre **un efect backwash al administrării asupra proiectării instrumentului de evaluare**. De exemplu, atunci când într-un examen nu avem posibilitatea furnizării testului fiecărui candidat și trebuie să apelăm la metoda transcrierii pe și de pe tablă a subiectelor; atunci calitatea testelor va fi redusă simțitor (nu putem utiliza: hărți, grafice, tabele, diagrame, texte literare de o anumită lungime, un număr suficient de itemi obiectivi ș.a.m.d.). Așa cum afirmam, fiecare elev trebuie să aibă propriu test de examen. Este numai o datorie morală, dar și o condiție științifică și pedagogică. Desenele, figurile, spațiile pentru răspunsuri, aranjarea specifică a itemilor în pagină - într-un cuvânt formatul prietenos al testului - constituie cerințe pedagogice importante. În plus, evitarea erorilor de scriere și rescriere a subiectelor, cât și eliminarea timpului necesar acestor operații au consecințe benefice asupra calității răspunsurilor care urmează să fie date de către elevi. Dacă din punct de vedere științific și psihopedagogic toată lumea este de acord cu necesitatea testelor de examen individuale, problemele care se invocă sunt de natură financiară și de păstrarea confidențialității/secretului testelor. În opinia noastră costurile sunt nesemnificative în raport cu alte costuri din învățământ (tipărirea manualelor, repararea și întreținerea școlilor ș.a). Costurile cresc pe măsură ce dorim ca testele să aibă calități tehnice superioare. O foarte bună validitate și fidelitate se obțin cu eforturi financiare sporite (de exemplu prin introducerea unor probe practice și experimentale). Este cunoscut faptul că un examen bun este costisitor. De aceea acesta trebuie proiectat în raport cu scopurile sale (pentru un examen cu miză mare se vor aloca resurse financiare mai mari). O parte din bani s-ar putea recupera din taxe modeste obținute de la candidați așa cum se întâmplă în alte sisteme de învățământ.

Principalele etape în administrarea testelor sunt:

(i) Tipărirea și distribuirea testelor:

- Realizarea variantelor finale pe computer.
- Tipărirea testelor pentru fiecare candidat.
- Pregătirea testelor pentru distribuire.
- Distribuirea testelor în condiții de securitate și păstrarea lor în centrele de examen în încăperi special amenajate.

(ii) Selectarea centrelor de examen:

- Colectarea datelor despre centre și candidații (număr, opțiuni pe discipline, ș.a).
- Construirea unor baze de date și actualizarea lor permanentă.

- Selecționarea centrelor.
- Păstrarea unei legături permanente cu centrele, în scopul actualizării informațiilor.

(iii) Selectarea și aprobarea comisiilor de examen și instruirea membrilor acestora (comisiile se organizează la nivel național, regional, și local - centre de examen):

- Selectarea examinatorilor șefi (președinți și vicepreședinți de comisii), corectorilor, supraveghetorilor și persoanelor însărcinate cu activități de secretariat.
- Elaborarea instrucțiunilor scrise pentru fiecare categorie de personal implicat în administrarea examenului (aceste instrucțiuni vor fi, în final, incluse în manualul de administrare al examenului).
- Formarea și informarea acestora.
- Monitorizarea procesului de desfășurare a examenului în centre, prin delegarea unor supervizori de la nivel central și județean.

(iv) Administrarea rezultatelor examenului:

- Rezolvarea contestațiilor.
- Transmiterea centralizată a datelor privind rezultatele candidaților.
- Întocmirea rapoartelor examenului la nivelul fiecărei comisii.

Așa cum remarcam, administrarea testelor este nu numai o activitate importantă, dar și complexă. De aceea coordonarea acestora trebuie realizată de către o instituție specializată, care dispune de expertiza și resursele necesare administrării standardizate și la standardele de calitate recunoscute pe plan internațional. Și mai trebuie făcută o remarcă. Întregul proces de administrare trebuie să fie **transparent** pentru public. Altfel este încurajată corupția sau, în cel mai bun caz, publicul are convingerea că aceasta există.

4.8 Diseminarea rezultatelor

În urma administrării testelor scrise se obțin o serie de date privind rezultatele elevilor. Uneori aceste date sunt completate cu alte informații provenite din aplicarea de chestionare, studierea documentelor școlare ș.a. Toate aceste date și informații vor fi introduse pe computer și preluate. Nu vom insista aici asupra modalităților de analiză statistică a datelor (statistică descriptivă, analiză de item, analiză factorială ș.a.), ci vom spune doar că toate concluziile desprinse în urma efectuării acestor analize trebuie publicate sub formă de rapoarte. Notele/mediile obținute la un examen reprezintă cea mai importantă etapă în desfășurarea sa din punctul de vedere al elevilor, părinților și publicului (Stoica A., 2000, p. 55-56). Situația este puțin diferită pentru o altă categorie de beneficiari ai rezultatelor - profesori și factori de decizie - în sensul că aceștia sunt interesați de rezultate nu numai din perspectiva notelor. Spre exemplu *feedback-ul* (rezultatele) unui examen național le oferă informații asupra: nivelului de pregătire al elevilor în raport cu standardele naționale, performanței elevilor în raport cu anumite variabile (sex, zone geografice etc.), procentului de ocupare a locurilor la diferite forme și profiluri de învățământ ș.a. Așadar, dacă rezultatele oferite primei categorii de beneficiari sunt cel mai bine exprimate prin note, prezentarea acestora pentru cea de-a doua categorie este mult mai utilă să se facă prin rapoarte asupra examenelor. Structura rapoartelor și tipul de informație furnizate depind deci de auditorul cărui acestea se adresează.

Astfel, există mai multe categorii de rapoarte:

- Rapoarte pentru factorii de decizie.
- Rapoarte pentru specialiști (cei care elaborează testele, curriculum, manuale școlare etc.).
- Rapoarte pentru profesori.
- Rapoarte pentru mass-media.

În elaborarea acestor rapoarte nu se pune problema tipului diferit de informații, ci a modalităților în care acestea sunt prezentate, în așa fel încât să fie utile și pe înțelesul celor cărora li se adresează.

Sumar de idei

Concepte-cheie:

- calități ale instrumentelor de evaluare (validitate, fidelitate, obiectivitate, aplicabilitate);

- teste elaborate de profesori - teste standardizate;
- teste criteriale - teste normative;
- teste de cunoștințe - teste de aptitudini;
- matrice de specificații; schema de notare;
- moderare internă și externă;
- bănci de itemi;
- pre-experimentarea și experimentarea itemilor;
- administrarea testelor;
- indicele de dificultate și indicele de discriminare.

Idei principale:

- Validitatea se referă la „faptul dacă testul măsoară ceea ce este destinat să măsoare.”
- Tipurile de validitate sunt: de conținut, de construct, concurentă, predictivă și de fațadă.
- Fidelitatea reprezintă calitatea unui test de a produce rezultate constante în cursul aplicării sale repetate.
- Aplicabilitatea reprezintă calitatea testului de a fi administrat și interpretat cu ușurință.
- Obiectivitatea reprezintă gradul de concordanță între aprecierile făcute de către evaluatori independenți în ceea ce privește un răspuns corect pentru fiecare din itemii testului.
- Principala diferență dintre testele criteriale și cele normative constă în modul în care sunt interpretate rezultatele candidaților. Primele comunică performanța absolută a unui elev, în timp ce testele normative comunică performanța relativă a unui candidat în raport cu ceilalți candidați.
- Matricele de specificații pot fi generale sau particulare (detaliate).
- Pe baza schemei de notare se corectează și se punctează testele elevilor.
- Notarea analitică are ca scop identificarea și analiza erorilor elevului.
- Notarea holistică (globală) este utilizată atunci când nu este necesar un *feedback* asupra erorilor.
- Banca de itemi este o colecție de itemi organizați, clasificați și catalogați ca și cărțile în bibliotecă, din care aceștia pot fi aleși cu scopul de a se crea teste.

Notă:

1. Exemplul adaptat după IGCSE, Distance Training for School-Based Assessment in Mathematics, University of Cambridge Local Examinations Syndicate, 1992, p. 28

Capitolul V. Tipologia itemilor

În demersul de proiectare a oricărei probe de evaluare, etapa imediat următoare definitivării matricei de specificații o constituie scrierea efectivă a itemilor, elementele componente de bază ale instrumentului de evaluare, pe baza obiectivelor de evaluare clar formulate.

Desigur, calitățile dorite ale testului ce urmează a fi produs vor depinde direct și de calitățile tehnice ale itemilor elaborați, nu întotdeauna transparente fără o analiză atentă și, uneori, metode statistice de analiză adecvate. Din acest motiv, considerăm că este necesară o privire mai amănunțită asupra diferitelor tipuri de itemi, a caracteristicilor acestora și a regulilor necesare de proiectare.

În cele ce urmează, vom opera cu următoarea definiție de lucru a **itemului**:

item = <întrebare> + <formatul acesteia> + <răspunsul așteptat>

Pentru proiectarea itemilor și utilizarea lor concretă, fiecare dintre aceste trei ele-mente este strict necesar, aducând un plus de informație util în formularea, analiza și revizuirea acestora. De asemenea, aceste elemente se află într-o strânsă interdependență. Astfel, fără precizarea clară a formatului în care intenționăm să proiectăm itemul, răspunsurile așteptate nu pot fi formulate, deci nici modalitățile de corectare și notare; de asemenea, analiza posibilelor răspunsuri așteptate poate aduce informații semnificative asupra clarității întrebării inițial formulate, precum și asupra gradului de adecvare a formatului acesteia.

Termenul reprezentat de **<răspunsul așteptat>** în definiția adoptată mai sus necesită o clarificare suplimentară. Prin acest răspuns așteptat înțelegem atât formularea unui răspuns „model” prezumat (în cazul itemilor în care acesta poate fi formulat în termeni absoluți - în special cazul itemilor obiectivi și semiobiectivi), cât și schema de corectare și notare a acestui răspuns, respectiv modul în care elemente ale răspunsului (sau criterii, în cazul itemilor cu răspuns deschis) vor fi evaluate și notate.

În teoria și practica evaluării există o gamă bogată de tipuri de itemi, care diferă prin anumite caracteristici ale acestora, itemi care se adecvează, fiecare, într-o mai mare sau mai mică măsură scopului instrumentului de evaluare și obiectivelor avute în vedere. Succesul actului evaluativ poate fi asigurat doar prin utilizarea acelor tipuri de itemi care, prin calitățile lor, satisfac cel mai bine contextul concret (determinat, desigur, de scopul testării, obiectivele de evaluare, comportamentele avute în vedere etc.).

Criteriile de clasificare a itemilor pot fi multiple, clasificările ce derivă din aplicarea acestor criterii fiind și ele numeroase.

Vom exemplifica, în cele ce urmează, două dintre cele mai frecvent utilizate. Astfel, având în vedere **tipul de comportament cognitiv solicitat pentru producerea răspunsului**, pot fi diferențiate două tipuri de itemi (Osterlind, 1998, p. 30):

(i) itemi care solicită **selectarea unui răspuns dintr-un** număr de variante oferite (*selected-responded items*);

- în acest caz, persoana evaluată primește răspunsul corect împreună cu variante (în general incorecte) la problema propusă, sarcina fiind aceea de a selecta răspunsul corect;
- sunt oferite în acest sens instrucțiuni clare privind modul în care se va selecta și înregistra răspunsul corect;

(ii) itemi care solicită **construirea, producerea** unui răspuns corect de către cel evaluat (*constructed-response items*);

- în acest caz, nu sunt oferite variante de răspunsuri, subiectul evaluat urmând să creeze un răspuns, o soluție la întrebarea/problema formulată;
- prin structura și formatul întrebării, răspunsul solicitat poate avea dimensiuni variabile, de la un cuvânt, o expresie numerică, o formulă etc., până la un eseu de dimensiuni relativ ample.

Unul dintre cele mai larg utilizate criterii de clasificare care a determinat și tipologia cea mai răspândită a itemilor este criteriul **obiectivității în corectare și notare**. Această obiectivitate în notare reprezintă mai ales măsura în care mai mulți corectori independenți cad de acord asupra răspunsului corect și îl punctează, deci măsura în care aceștia interpretează în același mod baremul de corectare și notare. În conformitate cu acest criteriu, o clasificare a itemilor este prezentată în tabelul următor, care include atât categoriile corespunzătoare criteriului amintit, cât și diferitele tipuri de itemi subsumate fiecăreia dintre acestea.

Categorii	Itemii obiectivi	Itemii semiobiectivi	Itemii subiectivi (solicitând răspuns deschis)
Tipuri de itemi	Itemii cu alegere duală Itemii cu alegere multiplă Itemii de tip pereche	Itemii cu răspuns scurt Itemii de completare Întrebări structurate	Rezolvare de probleme Eseu structurat Eseu liber (nestructurat)

Fiecare dintre aceste tipuri de itemi are caracteristici specifice, care determină gradul de adecvare și posibilitățile de utilizare în diferite contexte evaluative. Vom încerca, în cele ce urmează, să realizăm o privire comparativă asupra caracteristicilor acestor tipuri de itemi, pentru a evidenția elementele pe care este necesar să fie bazată selectarea și utilizarea unui anumit tip în proiectarea diferitelor instrumente de evaluare.

Proprietăți	Itemii obiectivi	Itemii semiobiectivi	Itemii subiectivi
Pot reprezenta un eșantion larg de obiective educaționale	++	+	--
Pot reprezenta un eșantion larg de conținuturi	++	+	--
Pot măsura abilitatea de a rezolva probleme noi	++	+	++
Pot măsura abilitatea de a organiza, integra sau sintetiza	--	+	++
Pot măsura gradul de originalitate sau demersul inovativ de rezolvare	--	+	++
Prezintă o fidelitate ridicată	++	+	--
Asigură o discriminare puternică	++	-	--
Asigură o potențială bază pentru diagnoză	++	+	--
Necesită un timp scurt pentru a fi produși	-	+	+
Necesită timp scurt pentru răspuns	++	+	--
Se corectează și notează rapid	++	-	--
Răspunsurile pot fi înregistrate electronic (cu ajutorul cititorului optic)	++	-	--
Permit evaluarea exprimării scrise	--	-	++
Încurajează originalitatea demersului și exprimării	--	-	++
Nu permit ghicirea răspunsului corect	-	+	+

După cum se poate observa din tabelul comparativ prezentat mai sus (preluat și adaptat după Walberg; Haertel, 1998, p. 448 și Gronlund, 1998, p. 44), fiecare dintre aceste tipuri de itemi are punctele sale tari și slabe în raport cu celelalte tipuri sau în raport cu anumiți factori (care pot influența sau chiar determina decizia utilizării lor într-un anumit moment și context). Semnificativă, este necesitatea unei bune cunoașteri a acestor caracteristici de către cei care proiectează probe de evaluare și atenta cântărire a calităților și limitelor fiecărui tip de item în momentul producerii și utilizării acestuia.

În continuare, vom analiza fiecare dintre categoriile și tipurile de itemi menționate, referindu-ne la **caracteristicile generale ale acestora, domeniile lor de utilizare, reguli specifice de proiectare, greșeli tipice și modalitățile de corectare și notare.**

5.1 Itemii obiectivi

Itemii obiectivi solicită din partea elevului evaluat selectarea răspunsului corect sau, uneori, a celui mai bun răspuns, dintr-o serie de variante existente și oferite.

Caracteristici generale

Itemii din această categorie sunt larg utilizați în practica evaluativă, datorită anumitor proprietăți caracteristice ale acestora:

- pot fi corecți și notați în mod obiectiv, în comparație cu toate celelalte tipuri de itemi;
- pot acoperi o plajă largă de obiective de evaluare și de elemente de conținut într-un interval de timp relativ scurt;
- sunt relativ ușor de administrat, corectat și notat;

- permit un *feedback* rapid și utilizarea informațiilor în scop diagnostic;
- există și posibilitatea înregistrării electronice a răspunsurilor (prin scanare, cu ajutorul cititorului optic) și analizarea lor rapidă, pornind de la aceste date;
- existând posibilități reduse de a interveni eventuale erori sau diferențe în corectarea realizată de persoane diferite sau la momente diferite în timp, prezintă o fidelitate ridicată.

Domenii de utilizare

- sunt adecvați evaluării în cadrul diferitelor discipline, în anumite domenii ale acestora, acest format având o mare flexibilitate;
- sunt adecvați măsurării unei game largi de obiective de evaluare, de la recunoașterea unor informații factuale, până la aplicarea corectă a unor principii și reguli în diferite contexte, familiare sau noi, sau de la simplul nivel al recunoașterii, până la cel al evaluării.

Reguli de proiectare

Având în vedere obiectivul principal avut în vedere de fiecare persoană care proiectează probe de evaluare, acela de a produce itemi și probe care pot oferi rezultate **valide și fidele**, este necesară respectarea unor reguli care pot conduce la crearea unor itemi având calitățile tehnice dorite. Este adevărat, însă, că în cazul itemilor obiectivi, timpul necesar producerii unor itemi având aceste calități este relativ mare, în comparație cu alte categorii. Având în vedere însă aspectele pozitive menționate mai sus, este limpede că trebuie atent cântărită decizia în favoarea itemilor obiectivi sau a altor forme.

Pentru fiecare dintre tipurile de itemi obiectivi posibil de proiectat, vom enunța un număr de reguli specifice, necesar a fi urmate în acest scop.

Modalități de corectare și notare

Itemii obiectivi au, prin definiție, un singur răspuns corect. Astfel, în cazul lor, corectarea se va realiza prin simpla acordare a punctajului în cazul obținerii acestui răspuns sau a neacordării nici unui punct, în cazul în care răspunsul așteptat nu a fost cel selectat. Nu operăm cu punctaje parțiale, având în vedere că nu există, în acest caz, răspunsuri parțial corecte. Deci, din punctul de vedere al corectării și notării, timpul necesar profesorului evaluator este sensibil mai mic decât în cazul altor itemi. De asemenea, obținând scorurile elevilor într-un timp foarte scurt, rezultatele pot fi imediat folosite în procesul educațional, în scop diagnostic sau în scopul în care proba de evaluare a fost proiectată. Din categoria itemilor obiectivi fac parte următoarele tipuri de itemi: **itemi cu alegere duală, itemi de asociere (de tip pereche) și itemi cu alegere multiplă.**

5.1.1 Itemi cu alegere duală

Sunt itemi pentru care elevul evaluat este solicitat să selecteze un răspuns corect din două răspunsuri posibile oferite.

Caracteristici generale

- se pot utiliza pentru evaluarea unei game destul de largi de obiective și de niveluri ale abilităților vizate, deși în practica evaluativă curentă sunt utilizați mai ales pentru evaluarea cunoștințelor factuale;
- pot îmbrăca formate diferite, dar care funcționează pe același principiu: selectarea unui răspuns de tipul: adevărat/fals, corect/greșit, da/nu;
- pot fi testate multe întrebări într-un timp relativ scurt, asigurând o bună acoperire a conținuturilor investigate;
- necesită un interval de timp relativ mic pentru răspuns;
- nu ocupă un spațiu foarte mare în economia testului;
- asigură o fidelitate mare, având un singur răspuns corect, care poate fi apreciat și notat fără riscul unor erori sau diferențe între corectori;
- există o șansă relativ mare ca elevul să ofere răspunsul corect prin ghicirea acestuia (50% șansă de ghicire), situație care poate fi atenuată prin utilizarea unui număr semnificativ de itemi de acest tip într-o probă de evaluare;

- nu permit utilizarea informațiilor obținute în scop diagnostic, simpla recunoaștere a faptului că o afirmație este falsă neaducând un plus de informație asupra gradului de stăpânire a obiectivului vizat.

Domenii de utilizare

- sunt utilizați pentru majoritatea disciplinelor, fiind relativ ușor de proiectat;
- sunt proiectați pentru contexte evaluative și domenii specifice de conținuturi în care pot fi formulate probleme care sunt fără echivoc adevărate sau false, corecte sau incorecte;
- pot acoperi o gamă largă de obiective și rezultate ale învățării, dintre care menționăm (Gronlund, 1998, p. 77):
 - generalizări într-un anumit domeniu;
 - comparații între concepte;
 - afirmații cauzale sau condiționale;
 - relații între două evenimente, concepte, principii;
 - explicații privind apariția unor fenomene;
 - explicații ale unor concepte și fenomene;
 - predicții asupra unor fenomene, evenimente;
 - etape într-o procedură/proces;
 - calcule numerice (aplicarea unor proceduri);
 - evaluarea unor afirmații referitoare la evenimente sau fenomene.

Reguli de proiectare

- itemii cu alegere duală este necesar să opereze strict cu fapte, situații, probleme dihotomice, având un singur răspuns corect dintre cele două oferite;
- este necesară testarea unui singur element de conținut/concept printr-un item (în caz contrar, valențele diagnostice ale rezultatelor obținute dispar, dar și adecvarea la obiectivul formulat);
- se va evita utilizarea unor propoziții lungi, acestea putând determina neclarități, îndepărtarea de la sensul real al sarcinii sau răspunsuri greșite cauza-te de dificultăți în lectura sarcinii; modul de exprimare va fi, de asemenea, pe cât de clar cu putință, lipsit de ambiguități;
- se va păstra un limbaj adecvat nivelului de vârstă al celor evaluați, pentru a nu distorsiona rezultatele obținute;
- se folosesc, pe cât posibil, propoziții afirmative în formularea sarcinii;
- instrucțiunile privind modul în care va fi selectat răspunsul trebuie să fie foarte clare.

Greșeli în proiectare

- În cazul în care, în proiectare, nu operăm cu fapte/contexte/probleme dihotomice (care permite fără dubiu variantele de răspuns corect sau incorect), subiectivitatea în corectare apare, precum și riscul corectării și notării în care intervin erori;
- folosirea negației, a dublei negații sau a altor formulări mai puțin clare, ridică neintenționat gradul de dificultate al itemului, îndepărtându-l în același timp de evaluarea obiectivului precizat.

Modalități de corectare și notare

- Având un singur răspuns corect, se corectează și notează prin acordarea punctajului doar cazul obținerii răspunsului așteptat.

Exemple:

Citește cu atenție afirmația următoare. Dacă o consideri adevărată, încercuiește litera **A**, iar dacă o consideri falsă, încercuiește litera **F**.

A F

Delta Dunării s-a format datorită mareelor.

Răspuns corect: **F**

A F

Expresia corectă a legii lui Ohm pentru o porțiune de circuit este dată de relația: $R = I/U$.

Răspuns corect: **F**

A F

Celulele gliale au rol în troficitatea mușchilor deoarece fac legătura între neuroni și vasele capilare.

Răspuns corect: **A**

Itemi cu alegere duală care solicită modificarea variantei false

Reprezintă o adaptare a itemilor cu alegere duală, având drept scop reducerea riscului de oferire a unor răspunsuri la întâmplare și ținând cont de principala critică adusă acestor itemi, aceea că recunoașterea unei afirmații false nu demonstrează și cunoașterea celei corecte.

Formatul principal este identic, instrucțiunea suplimentară care apare este aceea de a scrie afirmația corectă, în cazul în care aceea oferită este considerată falsă.

Exemplu:

Citește cu atenție afirmația următoare:

Dacă o consideri adevărată, încercuiește litera **A**, iar dacă o consideri falsă, încercuiește litera **F** și scrie, în spațiul rezervat, afirmația corectă.

A F

Temperaturile scăzute grăbesc circuitul apei în natură.

Răspuns corect: **F**

Afirmația corectă: *Temperaturile scăzute încetinesc circuitul apei în natură.*

5.1.2 Itemi cu alegere multiplă

Sunt itemi care solicită subiectului evaluat alegerea răspunsului corect sau a celui mai bun răspuns dintr-un număr de variante construite deja.

Sunt constituiți dintr-o **premisă** - reprezentând partea introductivă a itemului, stimulul pentru obținerea răspunsului, în care sarcina este formulată - și un număr de variante, răspunsurile posibile, dintre care unul corect, iar restul numiți **distractori**.

Caracteristici generale

- sunt itemi care permit obținerea răspunsurilor într-un interval de timp și un spațiu alocat rezonabile;
- acoperă un larg eșantion din aria conținuturilor într-un interval scurt de testare;
- sunt ușor de corectat;
- asigură o fidelitate ridicată;
- sunt flexibili, putând fi proiectați practic pentru orice nivel al abilităților și obiectivelor vizate;
- pot fi însoțiți de material-stimul narativ, grafic, cartografic etc.;
- oferă informații utilizabile în scop diagnostic, atunci când distractorii sunt construiți folosind greșeli tipice uzuale;
- rezultatele pot fi înregistrate electronic;
- excesiva utilizare a tehnicii poate avea efecte negative asupra învățării;
- rezultatele obținute pot fi influențate de ghicirea răspunsului corect;
- necesită un timp relativ lung de proiectare pentru a avea calitățile tehnice necesare obținerii unor informații valide și fidele.

Domenii de utilizare

- se pot proiecta itemi cu alegere multiplă practic pentru orice disciplină și domeniu evaluat;
- sunt adecvați evaluării oricărui nivel al abilităților și obiectivelor vizate, de la simpla recunoaștere a unor informații factuale, până la aplicarea unor algoritmi de calcul sau principii, identificarea unor exemple pentru concepte, fenomene, analiza unor situații noi, evaluarea unor contexte familiare sau noi;
- nu sunt adecvați evaluării capacității de exprimare în scris, de organizare a demersului sau a discursului pe o temă dată, de argumentare a unui punct de vedere sau a unei idei;

- utilizarea itemilor cu alegere multiplă cu mai mult de un răspuns corect face corectura mai dificilă și reduce obiectivitatea în notare.

Reguli de proiectare

• Premisa:

- trebuie să fie formulată în conformitate cu obiectivul de evaluare;
- trebuie formulată în termeni lipsiți de ambiguitate, care să nu ridice neintenționat gradul de dificultate al itemului;
- trebuie să conțină un verb, care să orienteze sarcina și să determine alegerea unui răspuns;
- va include cât mai mult din informația posibil repetitivă din variante, evitând lectura repetată neneesară a unor propoziții prea lungi;

• Distractorii:

- trebuie să fie plauzibili, altfel se reduce numărul de alternative din care este selectat răspunsul corect (și, implicit, probabilitatea de a ghici răspunsul crește);
- trebuie să fie independenți;
- nu trebuie să conțină indicii pentru alegerea răspunsului corect;
- nu pot fi construiți ca sinonime sau antonime ale unuia dintre ei;
- trebuie să fie construiți pe baza unor greșeli tipice ale subiecților, acest fapt sprijinind utilizarea răspunsurilor în scop diagnostic;
- instrucțiunile privind selectarea răspunsului corect și înregistrarea acestuia trebuie să fie clare;
- nu trebuie să existe ambiguități în variantele formulate, corectitudinea gramaticală fiind importantă;
- itemii cu alegere multiplă utilizați într-o probă de evaluare trebuie să fie independenți între ei, răspunsul la unul dintre aceștia neinfluențând (pozitiv sau negativ) răspunsul la ceilalți;
- utilizarea sintagmelor „toate cele de mai sus” sau „nici una dintre cele de mai sus” nu este recomandată în scrierea variantelor de răspuns;
- ordinea plasării răspunsurilor corecte trebuie să fie aleatoare; astfel, variantele vor fi ordonate logic: atunci când răspunsurile reprezintă un singur cuvânt vor fi așezate în ordine alfabetică, dacă sunt date numerice în ordine crescătoare, datele vor fi ordonate cronologic etc.

Greșeli în proiectare

- afirmațiile incorecte din punct de vedere gramatical (mai ales în cazul premisei sub forma unei afirmații incomplete) conduc la confuzii, având posibilitatea de a obține răspunsuri corecte/greșite indiferent de stăpânirea obiectivului de evaluare vizat;
- incoerența gramaticală existentă între afirmația incompletă (premisea) și unele dintre variante conduce la indicii neintenționate pentru respingerea unor variante pe alte baze decât judecarea în contextul obiectivului evaluat;
- utilizarea unor forme multiple de itemi cu alegere multiplă (propunând ca alternative de răspuns două sau mai multe dintre cele formulate) complică nejustificat itemul, evaluând mai mult abilitatea de a face față testării decât obiectivul;
- uneori sunt proiectați itemi de acest tip excesiv de lungi, care necesită un timp prea lung pentru a fi citiți, solicitând mai mult înțelegerea textului decât comportamentul evaluat;
- distractorii construiți uneori sunt evident greșiți, fiind eliminați din start;
- apar uneori indicii în premisă (cuvinte care se repetă în răspunsul corect), conducând la ghicirea răspunsului;
- distractorii au uneori semnificații apropiate (similari, sinonimici), eliminarea unuia conducând la eliminare mai multora;
- afirmațiile negative din premisă (sau cele care utilizează dubla negație) pot crea confuzie, înțelegere greșită a întrebării;

- varianta corectă mai lungă decât celelalte (construită de obicei astfel din dorința de corectitudine totală a răspunsului) determină alegerea acesteia automat, pe alte baze decât comportamentul care este măsurat.

Modalități de corectare și notare

- Itemii cu alegere multiplă având un singur răspuns corect sunt ușor de corectat, alocându-se punctajul doar pentru răspunsul așteptat.

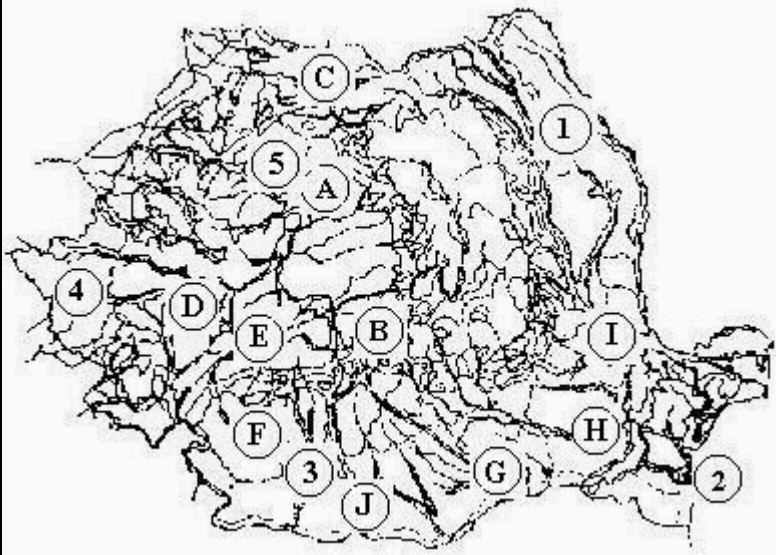
- Itemii având un număr mai mare de răspunsuri corecte sunt mai dificil de corectat și notat, crescând gradul de subiectivitate în acest proces; astfel, dacă avem două răspunsuri corecte, este necesar să fie luată o decizie asupra modului în care vor fi notate răspunsuri oferind un singur răspuns dintre cele corecte, sau unul corect și unul greșit. Este recomandabil, deci, utilizarea unui alt format de item în aceste cazuri (eventual transformarea itemului într-o serie de itemi de tip adevărat/fals).

Exemple:

Disciplina: Fizică Obiectiv de evaluare: elevii trebuie să recunoască unitatea de măsură asociată unei mărimi fizice Enunț: Încercuiți litera corespunzătoare răspunsului corect pentru afirmația de mai jos: Unitatea de măsură în S.I. pentru momentul forței este: } premise			
a. W x m	b. kg x m / s ²	c. J	d. Nx m
.....	
distractori			
Răspuns corect: d Disciplina: Geografie Obiectiv de evaluare: Elevii trebuie să identifice localizarea corectă a unui element geografic specific în raport cu un sistem de coordonate Enunț: Încercuiți litera corespunzătoare răspunsului corect pentru afirmația de mai jos: Față de centrul țării, situat la intersecția paralelei de 46° lat. N și a meridianului de 25° long. E, orașul Iași este situat în partea de:			
a. NE	b. NV	c. SE	d. SV
Răspuns corect: c			

- Itemii cu alegere multiplă pot fi formulați și pe baza unor materiale-stimul grafice, cartografice, imagini, fotografii, având în vedere specificul disciplinei în cadrul căreia vor fi utilizați.

- În funcție de materialul-stimul utilizat obiectivele de evaluare precizate vor putea avea o complexitate crescută, deci nu vor viza simpla recunoaștere a unor cunoștințe, fapte, date specifice domeniului evaluat.

Disciplina: Geografie	
Obiectiv de evaluare: Elevii trebuie să identifice elemente geografice specifice utilizând un suport cartografic	
Enunț: Studiați cu atenție harta. Încercuiți litera corespunzătoare răspunsului corect pentru fiecare dintre întrebările de mai jos:	
(1) Prin cifra 1 este reprezentat pe hartă orașul:	
a) Bacău	b) Iași
c) Piatra Neamț	d) Suceava
(2) Prin cifra 2 este reprezentat pe hartă orașul:	
a) Brăila	b) Constanța
c) Slobozia	d) Tulcea
	
(3) Prin litera A este reprezentată unitatea de relief denumită:	
a) Câmpia Transilvaniei;	
b) Podișul Târnavelor;	
c) Podișul Someșan;	
d) Subcarpații Transilvaniei;	
(4) Prin litera B este reprezentată grupa munților:	
a) Bucegi	b) Făgăraș
c) Parâng	d) Retezat
Răspunsuri corecte: (1) b; (2) b; (3) a; (4) b.	

5.1.3 Itemi de asociere (de tip pereche)

Reprezintă, de fapt, un caz particular de itemi cu alegere multiplă (în care mai mulți itemi împart un set comun de variante), fiecare dintre premise împreună cu setul răspunsurilor constituind, de fapt, un item cu alegere multiplă. Sunt formați, deci, din două liste și instrucțiuni care precizează modul în care se va realiza asocierea elementelor acestora și modul de înregistrare.

Domenii de utilizare

- pot fi utilizați pentru a măsura relațiile între diferite tipuri de informații factuale, dar nu numai: nume, date, evenimente, locuri, rezultate, termeni, expresii ale unor legi etc.
- sunt dificil de evaluat abilități de nivel înalt și obiective complexe cu acest tip de item.

Reguli de proiectare

- Este necesară furnizarea unor instrucțiuni foarte clare referitoare atât la tipul elementelor existente în fiecare coloană, cât și la spațiul și modul în care este solicitat răspunsul;
- Setul de premise, respectiv de răspunsuri trebuie să fie omogene;

- Numărul premiselor nu trebuie să fie prea mare, crescând astfel nejustificat nivelul de complexitate (se consideră, în general, că un număr de 4 - 5 premise este satisfăcător);
- Răspunsurile trebuie să fie relativ scurte (vor fi citite de mai multe ori pentru a rezolva itemul), ordonate logic;
- Numărul răspunsurilor trebuie să fie mai mare decât cel al premiselor, evitând corespondența directă între cele două seturi;
- Este necesară existența mai multor răspunsuri plauzibile pentru fiecare premisă (pentru a evita „ghicirea” răspunsului corect).

Greșeli în proiectare

- sunt similare celor enunțate pentru itemii cu alegere multiplă sau rezultă din regulile de proiectare.

Modalități de corectare și notare

- Răspunsurile corecte pentru fiecare dintre premise fiind unice, corectarea se va face dihotomic, prin alocarea punctajului în cazul răspunsului corect;
- În cazul obținerii unor răspunsuri (dintre care unele corecte, altele incorecte), este necesară luarea unei decizii privind modul de corectare și notare; această decizie trebuie să țină cont de scopul instrumentului de evaluare și de contextul în care evaluarea are loc, acest fapt permițând o notare „pozitivă” (punctaj alocat răspunsurilor corecte și zero puncte celor incorecte) sau „negativă” (mergând până la neacordarea nici unui punct în cazul existenței unor asocieri greșite).

Exemple:

Disciplina: Geografie Obiectiv de evaluare: Elevii trebuie să relaționeze localizarea corectă a unor elemente geografice specifice Enunț: Scrieți în spațiul liber din dreptul fiecărei cifre din prima coloană, cuprinzând numele unor râuri, litera corespunzătoare orașului situat pe fiecare râu, dintre cele din coloana a doua:	
Râuri	Orașe
____ 1. Bistrița ____ 2. Ialomița ____ 3. Jiu ____ 4. Mureș ____ 5. Someș	A. Arad B. Craiova C. Galați D. Giurgiu E. Iași F. Piatra Neamț G. Satu Mare H. Slatina I. Slobozia
Răspuns corect: F - 1; I - 2; B - 3; A - 4; G - 5.	

Disciplina: Fizică Obiectiv de evaluare: Elevii trebuie să recunoască unități de măsură asociate unor mărimi fizice. Enunț: Scrieți în spațiul liber din dreptul fiecărei cifre din prima coloană, cuprinzând numele unor mărimi fizice, litera corespunzătoare unității de măsură adecvate, dintre cele din coloana a doua:	
Mărimi fizice	Unități de măsură
____ 1. Moment cinetic ____ 2. Impuls mecanic ____ 3. Energie mecanică ____ 4. Accelerație	A. J B. J x s C. m / s ² D. kg / m ³ E. N / m F. N x s
Răspuns corect: B - 1; F - 2; A - 3; C - 4.	

5.2 Itemi semiobiectivi

Itemii semiobiectivi se îndepărtează, într-o oarecare măsură, de la maxima obiectivitate în corectare și notare asigurate de itemii obiectivi. Acești itemi solicită din partea elevului „producere” unui răspuns, de regulă scurt care va permite din partea profesorului evaluator sau examinator formularea

unei judecăți de valoare privind corectitudinea răspunsului oferit de subiect. Astfel, chiar dacă este vorba despre un răspuns așteptat reprezentat printr-un cuvânt, decizia privind influența scrierii corecte (ortografia) a răspunsului asupra scorului este necesar să fie luată. Din această categorie fac parte **itemii cu răspuns scurt, cei de completare și întrebările structurate**.

5.2.1 Itemi cu răspuns scurt

Reprezintă itemi care solicită producerea unui răspuns limitat ca spațiu, formă și conținut. Libertatea persoanei evaluate de a reorganiza informația primită și de a oferi răspunsul în forma dorită este foarte redusă, în același timp însă, este necesară demonstrarea abilității de a elabora și structura cel mai potrivit și scurt răspuns.

Caracteristici generale

- oferă posibilitatea de a evalua mai mult decât simpla recunoaștere și memorare a unor cunoștințe;
- solicită un anumit grad de coerență în elaborarea răspunsului;
- permit evaluarea unei game largi de cunoștințe, capacități și abilități;
- sarcina structurată și răspunsul scurt solicitat evită influența altor categorii de abilități decât cele vizate;
- pot sprijini discriminarea fină între fapte specifice;
- pot acoperi o arie largă de conținut;
- nu sunt vulnerabili față de răspunsurile furnizate la întâmplare, ghicite;
- necesită un timp relativ redus în proiectare;
- sunt relativ ușor de marcat și au un grad ridicat de obiectivitate, în condițiile elaborării unei scheme adecvate de corectare;
- pot exista variante de răspunsuri corecte, necesitând precizări clare în schema de corectare pentru acestea și pentru răspunsurile parțial corecte;
- răspunsul așteptat la întrebarea directă folosită trebuie oferit sub forma unei propoziții, a unui cuvânt, număr sau simbol;
- este necesar un număr relativ mare de itemi de acest tip pentru a acoperi o anumită arie de conținut;
- nu există riscul ghicirii răspunsului corect.

Domenii de utilizare

Pot fi menționate un număr de categorii de rezultate ale învățării pentru care poate fi utilizat acest tip de itemi (Gronlund, 1998, p. 96):

- cunoașterea terminologiei specifice;
- cunoașterea unor metode și proceduri;
- interpretarea unor date;
- aplicarea directă a unor legi, principii;
- abilitatea de a rezolva probleme numerice simple;
- abilitatea de a utiliza simboluri.

În același timp există și câteva limite în utilizare, dintre care menționăm:

- itemii cu răspuns scurt sunt mai puțin adecvați pentru evaluarea niveluri-lor cognitive superioare (analiza, sinteza, rezolvarea de probleme);
- răspunsul scurt solicitat nu este de natură să încurajeze dezvoltarea capacităților complexe.

Reguli de proiectare (Carey, 1988, p. 121):

- În cazul răspunsurilor numerice la itemi vizând rezolvarea unor probleme simple, este necesară precizarea tipului de răspuns așteptat și gradul de precizie al acestui răspuns.
- În cazul răspunsurilor numerice urmate de unități de măsură corespunzătoare, acestea trebuie clar precizate atât în întrebare, cât și după spațiul liber. Astfel sunt evitate eventualele neclarități ale întrebării și răspunsurilor oferite precum și corelarea clară cu obiectivul vizat.

- Este necesară evitarea unor întrebări mai puțin clare, care pot conduce la alte răspunsuri corecte decât cele vizate.
- Pentru itemii de completare este necesară evitarea excesului de spații albe, eliminându-se doar cuvinte-cheie.

Greșeli în proiectare

- nealocarea unui spațiu pentru răspuns poate determina formularea acestuia în termenii unui eseu și nu a unui răspuns scurt așteptat;
- formularea neclară (fără indicarea gradului de precizie cu care se dorește răspunsul, a unităților de măsură etc.) poate conduce la răspunsuri inadecvate sau incorecte;
- schema de corectare care nu prevede posibile răspunsuri alternative poate constitui o sursă de erori în corectare.

Modalități de corectare și notare

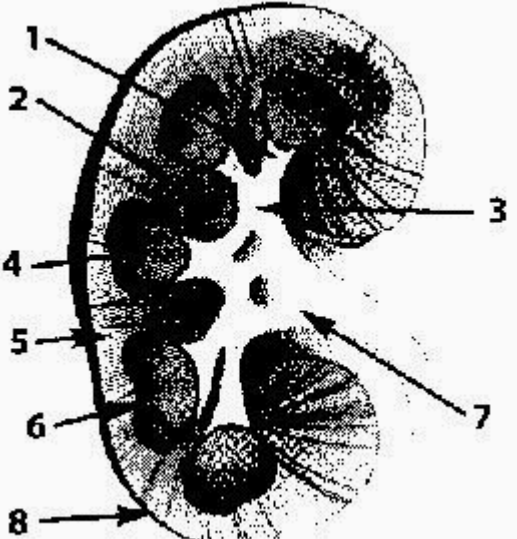
- Corectarea și notarea sunt mai dificile și necesită un interval mai lung de timp decât în cazul itemilor obiectivi, obținând uneori mai mult de un răspuns corect;
- Schema de corectare și notare trebuie elaborată cu multă atenție, necesitând în general și o etapă de pretestare, mai ales în cazul utilizării acestui tip de itemi în examinare.

Exemple:

Disciplina: Matematică
Obiectiv de evaluare: Elevii trebuie să rezolve exerciții simple utilizând operațiile matematice de bază
Enunț: Care este rezultatul efectuării următoarelor operații matematice? Scrie în spațiul punctat de mai jos fiecare rezultat corect.

1)	56 + 43	2)	78 - 25	3)	24 x 25
----	---------------------	----	---------------------	----	---------------------

Disciplina: Biologie
Obiectiv de evaluare: Elevii trebuie să recunoască elementele componente ale unor organe.
Temă: Scrieți în dreptul cifrelor de mai jos **denumirile corespunzătoare** ale componentelor rinichiului, vizibile în secțiunea longitudinală reprezentată alăturat:

	1. 2. 3. 4. 5. 6. 7. 8.
---	--

Răspuns corect:

1. calice mici	5. cortex renal
2. papilă renală	6. baza piramidei renale
3. calice mari	7. pelvis renal
4. piramide renale (Malpighi)	8. capsulă renală

5.2.2 Itemi de completare

O varietate a tehnicii răspunsului scurt o constituie **itemii de completare**, solicitând persoanei evaluate producerea unui răspuns care să completeze o afirmație incompletă sau un enunț lacunar și să îi confere valoare de adevăr. Diferența de formă a cerinței constă în faptul că în primul caz sarcina este formulată printr-o întrebare directă, în timp ce în al doilea caz este vorba despre o afirmație incompletă, răspunsul încadrându-se în contextul-suport oferit. Caracteristicile generale, domeniile de utilizare și modalitățile de corectare și notare sunt aceleași ca și pentru itemii cu răspuns scurt. În mod specific, intervin câteva reguli de proiectare.

Reguli de proiectare

- spațiile libere de dimensiuni variabile pot sugera, prin lungimea lor, răspunsul așteptat;
- excesul de spații albe în itemii de completare poate duce la pierderea sensului enunțului oferit spre analiză și completare, precum și la obținerea altor răspunsuri corecte decât cele așteptate;
- spațiul liber furnizat pentru producerea răspunsului trebuie să fie adecvat lungimii așteptate a acestuia, fără a da indicii asupra răspunsului corect. Astfel, în cazul mai multor itemi cu răspuns scurt succesivi, este recomandat ca spațiile pentru răspunsuri să aibă lungimi egale;
- în cazul itemilor de completare, nu este indicată utilizare drept suport a unor texte existente în manualele școlare. Consecințele nerespectării acestei cerințe ar fi încurajarea memorării, dar și testarea unor capacități și abilități diferite de cele vizate de fapt.

Exemplu:

Disciplina: Geografie

Obiectiv de evaluare: Elevii trebuie să identifice elemente geografice specifice conform proprietăților caracteristice enunțate ale acestora

Enunț: Completați afirmația de mai jos cu informația corectă și completă:

„Cel mai lung râu care curge integral pe teritoriul țării noastre este râul „

- De asemenea, itemii de completare se pot formula pe baza unui material grafic sau cartografic, în funcție de disciplină și de obiectivul de evaluare vizat.

Exemplu:

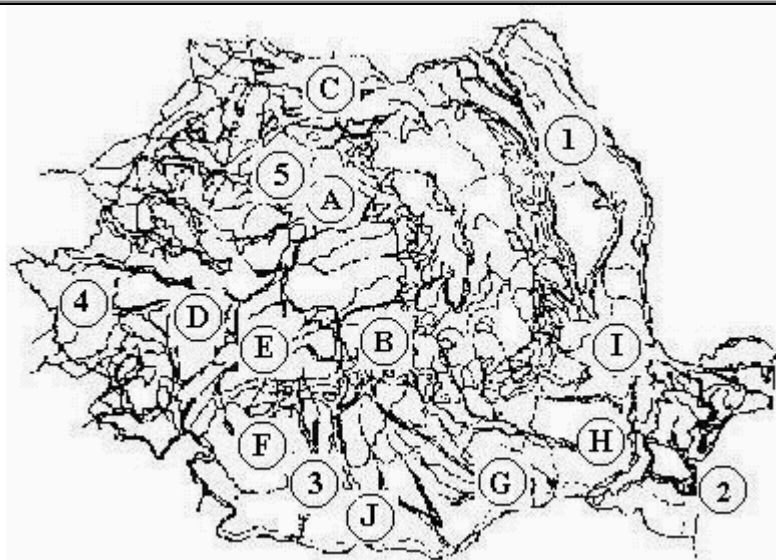
Disciplina: Geografie

Obiectiv de **evaluare**: Elevii trebuie să identifice corect localizarea unor elemente geografice specifice, utilizând un suport cartografic.

Enunț: Analizați cu atenție harta și completați spațiile libere din afirmațiile de mai jos cu informația corectă:

I. Orașele reprezentate pe hartă cu următoarele cifre se numesc:

- 1 4.....
2..... 5.....
3.....



II. Subdiviziunile unor unități majore de relief reprezentate pe hartă cu următoarele litere (A - E) se numesc:

- A D.....
B E.....
C.....

III. Râurile reprezentate pe hartă cu următoarele litere (F - J) se numesc:

- F I.....
G J.....
H.....

Răspunsuri corecte:

I. 1 - Iași; 2 - Constanța; 3 - Craiova; 4 - Timișoara; 5 - Cluj-Napoca;

II. A - Câmpia Transilvaniei; B - grupa (munții) Făgăraș; C - Depresiunea Maramureș; D - Munții Poiana Ruscă; E - Depresiunea Petroșani

III. F - Jiu; G - Argeș; H - Ialomița; I - Siret; J - Olt

5.2.3 Întrebări structurate

Întrebările structurate sunt sarcini formate din mai multe sub-întrebări, de tip obiectiv sau semiobiectiv, legate între ele printr-un element comun. Întrebările structurate acoperă spațiul liber aflat între tehnicile de evaluare cu răspuns liber și cele cu răspuns limitat impuse de itemii obiectivi. Acest tip de itemi oferă persoanei evaluate ghidare în elaborarea răspunsului, respectiv un cadru în care își realizează demersul. În elaborarea întrebărilor structurate, acestea pornesc în general de la un material-stimul (care poate fi reprezentat de diferite texte, date, diagrame, grafice, hărți etc.), urmat de un set de sub-întrebări, alte eventuale date suplimentare și un alt set de sub-întrebări.

Caracteristici generale

Având în vedere modul de prezentare și de structurare a sarcinii, întrebările structurate pot îndeplini două funcții importante:

- în primul rând oferă persoanei evaluate ghidare în elaborarea răspunsului,
- stimulează prin materialele utilizate producerea unui răspuns care se poate înscrie inclusiv în categorii vizând alte rezultate ale învățării decât reproducerea sau aplicarea;
- permit transformarea unui item de tip eseu într-o suită de itemi obiectivi, sau semiobiectivi, având ca principal efect creșterea fidelității în marcare;
- structurarea întrebărilor permite evaluarea unei game largi de cunoștințe, capacități și abilități, cu accent crescut pe nivelurile cognitive superioare;
- permit construirea progresivă a unei dificultăți și complexități dorite;
- permit crearea unui număr de sub-întrebări legate printr-o temă comună, ceea ce conduce la o abordare profundă a acelei teme, din diferite perspective;
- utilizează materiale-stimul auxiliare (hărți, grafice, diagrame etc.), ceea ce determină creșterea gradului de atractivitate al acestui tip de itemi, dar și posibilitatea evaluării unor capacități specifice, dificil de evaluat în alt context;
- stimulează atât dezvoltarea capacităților cu un nivel mai ridicat de complexitate, cât și originalitatea și creativitatea celui evaluat;
- materialele auxiliare având calitatea tehnică necesară sunt relativ dificil de proiectat;
- răspunsul la o sub-întrebare depinde uneori de cel oferit la sub-întrebări precedente, necesitând indicații clare în schema de marcare;
- costurile sunt relativ ridicate și timpul necesar proiectării întrebărilor structurate este mai lung.

Reguli de proiectare

- sub-întrebările vor fi în general cu răspuns deschis, dar solicitând produce-rea unui răspuns relativ scurt;
- gradul de dificultate al sub-întrebărilor va crește progresiv spre sfârșitul itemului; gradul de dificultate al cerinței poate fi asociat uneori cu solicitarea unui răspuns mai amplu;
- fiecare sub-întrebare poate să fie autoconținută și să nu depindă de răspunsul corect la întrebarea precedentă;
- în cazul întrebărilor structurate progresive, în care este continuată o etapă prin cerința următoare, sunt necesare indicații clare în schema de corectare, pentru a evita dubla depunere pentru o singură greșeală;
- spațiul liber oferit pe foaia de răspuns trebuie să fie în concordanță cu lungimea reală a răspunsului așteptat;
- sub-întrebările trebuie să fie în concordanță și strictă corelație cu materialele-stimul utilizate;
- materialele auxiliare utilizate trebuie să aibă calitatea tehnică adecvată su-portului necesar, fără a ridica dificultăți suplimentare sarcinii;
- schema de marcare trebuie să fie proiectată în același timp cu formularea întrebării.

Gradul de dificultate al întrebărilor structurate este conferit de formatul acestora, de materialele utilizate și de tipul sarcinii, respectiv de abilitățile, competențele care sunt vizate. Modificând unul sau mai multe dintre aceste elemente, se pot obține întrebări structurate pe diferite niveluri de dificultate, adecvate unor scopuri și populații diverse.

Caracteristicile întrebărilor structurate prezentate mai sus le recomandă pentru utilizarea în cadrul oricărei discipline, cu atenția necesară acordată limitelor amintite. În cazul unei proiectări atente și imaginative, pot sprijini cu succes evaluarea unor demersuri cu adevărat originale și creative, analiza unor situații noi și explicarea/rezolvarea acestora prin modalități exersate sau noi. Schema de corectare și notare este punctul egal important al proiectării și utilizării acestor itemi, de calitatea acestora fiind direct legată eficiența aplicării lor.

Exemplu (preluat din Examenul real de Bacalaureat 2000, Proba de Istoria românilor):

Disciplina: Istoria românilor

Obiectiv de evaluare: Elevii trebuie să demonstreze capacitatea de a selecta informații istorice din surse istorice date;

Elevii trebuie să fie capabili să stabilească relații de tip cauză - efect.

Enunț: Citiți cu atenție textul de mai jos:

”Creștinismul daco-romanilor ca și limba lor este de caracter latin și adoptarea lui a avut loc treptat, prin difuzarea în masa locuitorilor, nicidecum printr-o hotărâre a unei autorități centrale (rege, principe) ca și în alte părți.”

(Dinu C. Giurescu, despre adoptarea creștinismului)

Pornind de la acest text, răspundeți următoarelor cerințe:

1. Transcrieți modul de adoptare treptată a creștinismului de către daco-romani. **2 puncte**
2. Menționați o dovadă a răspândirii creștinismului în limba latină. **2 puncte**
3. Numiți un alt popor care a adoptat creștinismul. **3 puncte**
4. Explicați, în maximum 10 rânduri, un motiv pentru care răspândirea creștinismului în nordul Dunării nu a fost impusă de o autoritate centrală. **5 puncte**

Schema de corectare și notare

1. **2 puncte** pentru răspunsul: „ prin difuzarea în masa locuitorilor”.
2. **2 puncte** pentru oricare dovadă a răspândirii creștinismului în limba latină.
3. **3 puncte** pentru menționarea oricărui alt popor creștin.
4. **2 puncte** pentru menționarea oricărui motiv, corect din punct de vedere istoric.
2 puncte pentru oricare explicație, corectă din punct de vedere istoric.
1 punct pentru respectarea limitei de spațiu.

Total 12 puncte

Disciplina: Geografie

Obiective de evaluare: Elevii trebuie să fie capabili:

1. Să analizeze date cantitative prezentate în diferite formate;
2. Să stabilească ordonări după date cantitative oferite sau calculate;
3. Să identifice relații între fenomene, procese, date geografice specifice;
4. Să formuleze concluzii/extrapolări pe baza unor date oferite.

Temă:

I. În tabelul următor sunt prezentate evoluțiile producției de grâu în trei ani diferiți (1980, 1990 și 1998), pentru trei țări având cele mai mari producții.

Statul	Producția de grâu (mil. t.)			Creșteri 1998/1980	
	1980	1990	1998	mil. t.	%
China	55,1	98,2	110,0	54,9	199
S.U.A.	64,2	74,4	69,4	5,2	108
India	32,2	49,8	66,0	33,8	204

Analizați acest tabel și răspundeți, pe baza lui, la următoarele întrebări:

1. Precizați ordinea descrescătoare a acestor țări, după producția de grâu în anii 1980 și 1998:
1980
1998

2. Precizați ordinea țărilor după creșterile proporționale în acest interval de timp:
.....

II. Comparați tabelul de mai sus cu următoarele date privind evoluția producției mondiale de grâu pentru aceiași ani și populația totală.

	Anii		
	1980	1990	1998
Producția de grâu (mil. t.)	495,0	592,9	590,1
Populația (mild. loc.)	4,5	5,3	6,0

1. Completați tabelul următor:

Producția de grâu	Pondere din total (%) mondial	
	1980	1998
China		18,6

S.U.A.		11,7	
India		11,2	
2. Realizați un scurt comentariu pe baza datelor obținute:.....			
III. Urmăriți în continuare datele din tabelul următor, care redă creșterea populației acestor țări în aceeași perioadă de timp (1980-1998), în valori relative (%):			
China		126%	
S.U.A.		120%	
India		144%	
1. Comparați creșterea relativă a producției de grâu și a populației acestor țări, stabilind o ordine a țărilor după ambele criterii: (1) (2) (3) 2. Motivați, pe scurt, această ordine.			
Răspuns corect:			
I	1. 1980: S.U.A., China, India; 1998: China, S.U.A., India		
	2. India, China, S.U.A.		
II	1. China 11%, S.U.A. 13%, India 6%		
	2. Exemplu: Dacă ponderea S.U.A. în totalul producției de grâu a scăzut foarte puțin (de la 13% în 1980 la 11,7% în 1998), ponderea celorlalte două țări a crescut semnificativ (India aproape și-a dublat această valoare).		
III	1. India, China, S.U.A.		
	2. India a înregistrat cele mai mari creșteri proporționale atât la populație (144%), cât și la producția de grâu (204%), în același interval de timp; China a înregistrat creșteri mari ale producției de grâu (199%) și creșteri relative mijlocii ale populației (120%); S.U.A. a avut creșteri aproape comparabile ale populației și ale producției de grâu.		

5.3 Itemi subiectivi (care solicită un răspuns deschis)

Itemii subiectivi solicită din partea subiectului evaluat un răspuns, care este foarte puțin orientat în elaborarea sa prin structura sarcinii sau care nu este deloc orientat. Astfel, respondentul va decide singur care sunt elementele pe care le va include în răspunsul său și care nu, ce lungime va avea acest răspuns (în unele cazuri). Sunt numiți și itemi „subiectivi”, datorită situării la polul opus itemilor obiectivi din punct de vedere al obiectivității în corectare și notare; în alte sensuri, această denumire ar putea părea oarecum improprie, subiectivitatea neintervenind nici în formularea sarcinii, nici în modul concret de corectare, schema de corectate și notare încercând să atenueze acest posibil risc și să crească nivelul fidelității.

Caracteristici generale

- reprezintă o categorie de itemi necesar a fi utilizați numai atunci când se doresc evaluate rezultate complexe ale învățării, abilități de tip analiză, argumentare, sinteză, organizare a demersului sau a discursului și când nu putem utiliza în același scop alt format de itemi;
- numărul itemilor cu răspuns deschis posibil de inclus într-un test este relativ mic, din cauza timpului lung necesar pentru construirea răspunsurilor; deci, gradul de acoperire a domeniului evaluat nu este foarte mare;
- corectarea și notarea nu prezintă un grad mare de obiectivitate, existând în general variații între scorurile acordate de diferiți corectori;
- de exemplu, un singur eseu este considerat insuficient pentru a oferi o măsură acceptabilă a abilității de a scrie sau de a susține un argument;
- sunt relativ simpli de proiectat, necesitând în schimb un timp mai îndelungat pentru corectare și notare; din acest punct de vedere, informațiile obținute din analiza rezultatelor nu pot fi utilizate imediat în procesul educațional.

Domenii de utilizare

- sunt proiectați și utilizați pentru obiective și situații de evaluare în care interesează în mod deosebit demersul subiectului în producerea unui răspuns, nu întotdeauna unul singur posibil sau corect;
- măsoară abilități de nivel înalt, rezultate complexe ale învățării, în cadrul majorității disciplinelor;
- sunt cei mai adecvați pentru măsurarea abilității de exprimare în situații în care nu există limită de spațiu și răspunsul se poate formula în modul ales de respondent.

Reguli de proiectare

- întrebarea formulată trebuie să fie adecvată obiectivului de evaluare vizat, adecvată de asemenea populației căreia i se adresează (interesantă, capabilă să stimuleze un răspuns cât mai apropiat de adevărata măsură a abilităților vizate ale subiecților);
- în măsura în care este posibil și se adecvează obiectivului de evaluare, este preferată utilizarea itemilor care oferă o orientare a răspunsului subiectului, în termeni de dimensiuni ale răspunsului așteptat, criterii de evaluare ale acestui răspuns.

Modalități de corectare și notare

Având în vedere principalele scopuri ale etapei corectării și notării - reducerea erorii în apreciere dar și posibilitatea extinderii informației rezultate din evaluare pentru îmbunătățirea *feedback-ului* - modalitățile de corectare și notare pot îmbrăca două forme:

- **analitic:** constă în agregarea punctelor acordate pentru diferite elemente discrete ale răspunsului așteptat;
- **holistic:** constă în formularea unei judecăți globale asupra răspunsului acordat, în conformitate cu anumite criterii de evaluare sau descriptori de performanță, sau prin simpla încadrare într-o anumită categorie prin compararea răspunsurilor între ele.

5.3.1 Rezolvare de probleme

Rezolvarea de probleme sau, mai corect spus de situații problemă, reprezintă un tip de sarcină în cadrul căreia elevul este confruntat, în cele mai multe cazuri, cu o situație nouă, nefamiliară pentru care nu este evidentă o soluție predeterminată. În acest sens este necesar să accentuăm faptul că o adevărată situație - problemă nu are o soluție imediată, fără a parcurge un număr de etape, dintre care menționăm:

- identificarea problemei;
- analiza și selectarea datelor relevante;
- formularea unor ipoteze care urmează a fi evaluate și, eventual, validate;
- identificarea metodei/procedurii de rezolvare (în unele cazuri, identificarea unor metode diferite care permit comparații);
- crearea unor modele;
- propunerea unei soluții;
- evaluarea soluției obținute;
- formularea concluziei, implicând și eventuale considerații asupra acurateței soluției obținute sau a posibilelor surse de erori.

Desigur, nu toate sarcinile de tip rezolvare de probleme permit parcurgerea tuturor acestor etape, însă acestea reprezintă un potențial important care poate fi utilizat în mod divers.

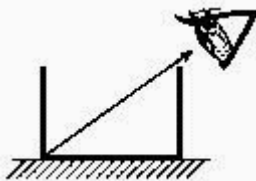
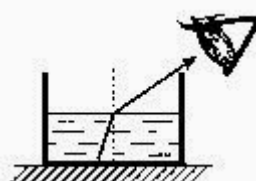
Există o mare diversitate de situații-problemă, care se dezvoltă între două extreme:

- situații „închise”, în care sunt oferite majoritatea datelor necesare, scopul este clar specificat, iar strategia de rezolvare este sugerată prin succesiunea cerințelor;
- situații complet „deschise”, în care sunt sugerate doar anumite elemente pentru a găsi soluția, în sarcina persoanei evaluate rămânând alegerea strategiei dintr-o gamă largă de posibilități și planificarea completă a demersului.

Caracteristici generale

- principalul avantaj al utilizării îl reprezintă stimularea gândirii creative;
- sprijină transferul în interiorul unui domeniu sau al mai multor domenii;
- permit utilizarea unor materiale - suport variate, cu o puternică reflectare în viața cotidiană;
- permit selectarea unei metode din mai multe posibile, având în vedere diferite criterii de selecție;
- se pot estima și analiza posibile erori;
- solicită atenție în proiectare și un timp relativ lung pentru corectare și notare;
- este important a se evita transformarea rezolvării de probleme în demonstrarea unei rutine de rezolvare; în acest sens diversificarea sarcinilor joacă un rol important;
- formularea cerințelor trebuie să fie în perfectă concordanță cu obiectivul de evaluare avut în vedere și cu tipul de situație - problemă care a fost ales.

Exemplu (preluat din Examenul real de Bacalaureat 2000, Proba de Fizică, Profilul real și tehnologic):

Disciplina: Fizică		
Obiectiv de evaluare: Elevii trebuie să aplice legi, principii ale fizicii, în rezolvarea unor probleme întâlnite în viața cotidiană.		
	Din poziția în care se află ochiul unui observator poate fi văzut în întregime peretele vertical al vasului din figură, dar nu poate fi văzut fundul acestuia. Dacă în vas se toarnă o cantitate de apă, se va putea vedea o parte din fundul vasului sau, dimpotrivă, nu se va mai vedea în întregime peretele vertical? Justificați răspunsul. 8 puncte	
Schema de corectare și notare: Efectuarea unui desen corect în care să fie ilustrat mersul razei de lumină care delimitează domeniul de observare 8 p din care:		
	- pentru considerarea refracției la suprafața de separare apă-aer	2p
	- pentru considerarea legii refracției	2p
	- r (unghiul de refracție) $>$ i (unghiul de incidență) pentru situația dată	2p
	Concluzia este că o parte din fundul vasului devine observabilă.	2p
Total :		8 puncte

5.3.2 Eseul

Reprezintă un tip de item în care subiectul evaluat este solicitat să producă un răspuns liber, în conformitate cu anumite cerințe sau criterii formulate.

În funcție de dimensiunea răspunsului așteptat, pot fi proiectate două categorii de eseuri:

- **eseul cu răspuns scurt** (minieseul), în care limita de spațiu este clar precizată în cadrul cerinței;
- **eseul cu răspuns extins**, în care dimensiunea răspunsului este la alegerea respondentului, singura limită care acționează fiind cea a timpului avut la dispoziție.

În funcție de gradul de detaliere a cerințelor și criteriilor formulate pentru elaborarea răspunsului, deosebim:

- **eseul structurat**, în care răspunsul așteptat este orientat prin cerințele formulate, libertatea de organizare a răspunsului nefiind totală;
- **eseul liber**, în care nu sunt creionate jaloane pentru răspunsul așteptat, la alegerea persoanei evaluate fiind atât elementele care vor fi incluse în răspuns, cât și modul lor de organizare; această a doua categorie răspunde mai ales unor obiective vizând originalitatea, creativitatea, scrierea imaginativă.

Caracteristici generale

- eseurile pot acoperi o gamă largă de obiective, de la cunoaștere, până la analiză și evaluare;
- având în vedere lungimea răspunsului așteptat, mai ales în cazul eseurilor extinse, timpul necesar pentru producerea răspunsului este relativ mare, deci pot acoperi o arie relativ mică de conținuturi/teme abordate într-un interval de testare;
- timpul necesar pentru proiectarea acestui tip de item este relativ mic, însă crește semnificativ timpul necesar corectării și notării;
- având în vedere îndepărtarea de „capătul obiectiv” al continuumului tipurilor de itemi, crește în mod semnificativ subiectivitatea în notare, fapt care duce la

scăderea fidelității; în acest sens, schema de corectare și mai ales selectarea modalității de corectare joacă un rol esențial.

Eseu structurat

Exemplu (preluat din *Modele de teste și bareme*, S.N.E.E., Editura ProGnosis, 2000):

Disciplina: Limba și literatura română

Obiectul de evaluare: Elevii trebuie să elaboreze un eseu structurat pe o temă dată, urmărind îmbinarea diferitelor categorii de cerințe formulate.

Scrie o compunere/eseu, prin care să înfățișezi statutul social și psihologic al personajului Costache Giurgiuveanu, din romanul *Enigma Otiliei*, de C. Călinescu. În realizarea lucrării, vei avea în vedere următoarele:

- prezentarea pe scurt a formulei estetice folosite de romancier în construcția personajelor;
- identificarea tipului uman reprezentat de Costache Giurgiuveanu;
- argumentarea, pornind de la întâmplări și situații selectate din cuprinsul romanului, a tipologiei identificate;
- prezentarea raporturilor dintre Costache Giurgiuveanu și cel puțin două dintre personajele romanului.

Atenție! Se recomandă ca eseuul tău să nu depășească 2-3 pagini. Ordinea integrării cerințelor în cuprinsul eseului este la alegerea ta. Compararea cu mai mult de două personaje din roman este apreciată cu până la patru puncte în plus la nota finală (câte două puncte pentru fiecare personaj)

.....**40 puncte (36 p. + 4 p.)**

Schema de corectare și notare:

1. Metoda balzaciană, dar, cum spunea Nicolae Manolescu, „un balzacianism fără Balzac”, polemic și critic; importanța „decorului” în conturarea trăsăturilor personajelor, narațiune la persoana a III-a etc.	4 p.
2. Avarul	4 p.
3. Argumentarea tipologiei identificate cu ajutorul unor întâmplări și situații semnificative	4 p.
- argumentarea nesigură, bazată pe întâmplări și situații nu întotdeauna relevante	2 p.
4. Prezentarea raporturilor dintre Costache Giurgiuveanu și două personaje ale romanului	(2 p. + 2 p.) = 4 p.
- prezentarea aproximativă/nesigură/parțială	(1 p. + 1 p.) = 2 p.
5. Compararea/paralela dintre Costache Giurgiuveanu și alte două personaje ale romanului	(2 p. + 2 p.) = 4 p.
- prezentare vagă, elemente ale portretului-tip ne semnificative:	(1 p. + 1 p.) = 2 p.
6. Organizarea ideilor în scris	4 p.
- text clar organizat, cu echilibru între cele trei componente, introducere-cuprins-încheiere; construcția paragrafelor subliniază ideile; ideile principale și cele secundare au o pondere echilibrată	4 p.
- părțile componente ale textului pot fi ușor recunoscute, dar nu există un echilibru între ele; ideile sunt, în general, subliniate prin paragrafe	2 p.
- plan vag de structurare a textului, în care părțile componente ale textului sunt insuficient marcate; trecerea de la o idee la alta nu este marcată în nici un fel	1 p.
7. Abilități analitice și critice	4 p.
- o foarte bună relație idee - argument, coerența construcției; succesiunea logică (motivată) de la o idee la alta, ceea ce face textul, în totalitatea sa, convingător	4 p.
- o evidentă încercare de a susține ideile prin argumente, chiar dacă rezultatul nu este întotdeauna convingător	2 p.
- afirmații rareori susținute de argumente; argumente nerelevante	1 p.
8. Folosirea limbii	4 p.

- vocabular adecvat situației, sigur ales și variat	4 p.
- vocabular adaptat susținerii ideilor, dar cu ezitări în selectarea cuvintelor	2 p.
- vocabular restrâns, monotonic, repetiții de cuvinte, impresie generală de stângăcie	1 p.
9. Ortografie	3 p.
- 0-1 erori	3 p.
- 2 erori	2 p.
- 3 erori	1 p.
10. Punctuație	4 p.
- 0-1 erori	4 p.
- 2 erori	3 p.
- 3 erori	2 p.
- 4-5 erori	2 p.
11. Așezarea în pagină	1 p.
Total:	40 p.

5.3.3 Eseu liber (nestructurat)

Exemplu

Disciplina: Geografie

Obiectiv de evaluare: Elevii trebuie să fie capabili să explice relația dintre elemente componente ale unor procese geografice specifice.

Explicați repartitia precipitațiilor medii anuale pe întinderea teritoriului țării noastre.

Schema de corectare și notare: Se acordă punctajul maxim (**6 p.**), câte **2 p.** pentru precizarea a oricare trei caracteristici (elemente explicative) dintre cele de mai jos:

- cantitatea medie anuală de precipitații este de 640 mm.
- precipitațiile sunt răspândite neuniform (cel mai important factor fiind relieful)
- cu creșterea altitudinii crește și cantitatea de ploi
- în munți plouă cel mai mult (1000 mm. anual)
- la câmpie plouă destul de puțin (sub 500 mm.)
- precipitațiile nu cad uniform în cursul anului
- luna cea mai ploioasă este iunie, iar cea mai uscată februarie
- iarna cad precipitații sub formă de zăpadă (la cele mai mari înălțimi zăpada se menține la peste 150 zile/an)

Total 3 x 2p . = 6 p.

Sumar de idei

Concepte-cheie:

- item;
- itemi obiectivi;
- itemi semiobiectivi;
- itemi subiectivi;
- itemi cu alegere duală;
- itemi cu alegere multiplă;
- itemi de asociere (de tip pereche);
- itemi cu răspuns scurt;
- itemi de completare;
- întrebări structurate;
- itemi de tip rezolvare de probleme;
- eseu structurat;
- eseu liber;

- schemă de corectare și notare;
- barem de corectare și notare;
- premisă;
- distractori

Idei principale:

• Itemii reprezintă elementele componente de bază ale instrumentelor de evaluare, fiind proiectați în conformitate cu obiectivele de evaluare formulate.

• **Item = <întrebare> + <formularul acesteia> + <răspuns așteptat>**

- Având în vedere criteriul obiectivității în corectare și notare, itemii se pot clasifica în itemi obiectivi, itemi semiobiectivi și itemi subiectivi.
- Itemii obiectivi solicită selectarea răspunsului corect dintr-o serie de variante oferite.
- Itemii obiectivi sunt itemi caracterizați printr-o fidelitate ridicată, permițând în același timp testarea unei game largi de obiective de evaluare și elemente de conținut într-un timp relativ scurt.
- În categoria itemilor obiectivi se înscriu: itemii cu alegere duală, itemii cu alegere multiplă, itemii de asociere (de tip pereche)
- Itemii semiobiectivi solicită elevului construirea unui răspuns, de mici dimensiuni, permițând evaluarea unei game largi de obiective și utilizarea unor materiale auxiliare variate.
- În categoria itemilor semiobiectivi sunt incluși: itemii cu răspuns scurt, itemii de completare și întrebările structurate.
- Itemii subiectivi (care solicită un răspuns deschis) sunt utilizați pentru evaluarea unor obiective de complexitate ridicată, solicitând decizia personală a subiectului privind demersul, modul de organizare a ideilor și, uneori, originalitatea abordării.
- Itemii subiectivi solicită un timp mai lung pentru corectare și marcare, bareme foarte clare și detaliate și proceduri specifice de moderare a corectorilor pentru a obține rezultate fidele
- În categoria itemilor subiectivi se înscriu rezolvarea de probleme, eseul structurat și eseul liber.

Capitolul VI. Reforma evaluării rezultatelor școlare în România

În ultimii ani, în contextul general al reformei învățământului românesc, o reformă coerentă și profundă a fost introdusă la toate nivelurile învățământului preuniversitar. Schimbările de esență au vizat atât evaluarea curentă, cât și evaluările externe (examele și testările naționale). Astfel, în domeniul evaluării la clasă a fost elaborat și aplicat cu succes un nou sistem de apreciere prin calificative în ciclul primar și a fost îmbunătățită evaluarea rezultatelor elevilor în gimnaziu și liceu, în strânsă legătură cu metodele moderne de evaluare, noua structură a anului școlar (perioada de recapitulare - evaluare) și noul curriculum. În ceea ce privește examenele, a fost proiectat și s-a desfășurat pentru prima dată examenul național de Capacitate, iar rezultatele acestuia au constituit baza admiterii în licee și școli profesionale. Bacalaureatul a fost complet reformat începând cu modul de formulare a subiectelor și încheind cu administrarea sa, astfel încât acesta a devenit un examen din ce în ce mai credibil pentru a constitui unul din criteriile de admitere în învățământul superior. Nu în ultimul rând, a fost proiectat un sistem de evaluări naționale cu scop diagnostic (pe eșantioane reprezentative de elevi), din care s-au realizat, până în prezent, câteva testări semnificative la clasa a IV-a.

6.1 Evaluarea curentă: necesitate, caracteristici și modalități de realizare

Evaluarea curentă (vezi și Stoica A., 2000, p. 35 și următoarele) este parte integrantă a procesului de instruire. Aceasta se desfășoară continuu pe parcursul procesului de instruire, prilej cu care profesorul obține informațiile necesare privind progresul școlar, fiind astfel în măsură să regleze metodele de predare și să orienteze activitățile de evaluare. *Feedback-ul* furnizat de evaluarea curentă este util, în egală măsură și elevului, prin evidențierea stadiului la care acesta a ajuns în atingerea obiectivelor educaționale. Evaluarea curentă - în special componenta sa formativă - are, din acest punct de vedere, rolul de a apropia pozițiile principalilor actori ai educației: profesorul și elevul. În mod normal ambii sunt interesați în a găsi căile, metodele și mijloacele de a atinge obiectivele instruirii la nivelul dorit.

Principala caracteristică a evaluării curente este posibilitatea utilizării tuturor metodelor și instrumentelor de evaluare pe care profesorul le are la dispoziție. Fie că este vorba de metodele tradiționale sau de cele complementare, profesorul le alege pe cele mai adecvate disciplinei de învățământ, obiectivelor instruirii, tipului de conținut și particularităților de vârstă ale elevilor. În acest fel, fiecare elev are șansa să demonstreze priceperi și capacități, care nu pot fi puse în evidență în timpul limitat oferit de un examen. Pe de altă parte, unele metode și instrumente necesare pentru a evalua cunoștințele și capacitățile elevului sunt complexe, dificil de proiectat și supuse erorilor. Aceste erori nu pot fi eliminate, dar pot fi micșorate până la o limită acceptabilă, numai prin planificarea riguroasă a formelor, metodelor și tehnicilor de evaluare, precum și prin utilizarea lor complementară.

Evaluarea se realizează diferit de la un profesor la altul. Este evident că acțiunile evaluative curente nu pot fi standardizate - și nici nu este de dorit acest lucru - dar acestea trebuie să se bazeze pe obiectivele educaționale. În programele de studiu se pot sugera diferite metode și instrumente de evaluare, ca și sarcini specifice fiecărei discipline. Spre exemplu, se pot enunța teme pentru investigații în clasă, proiecte și elemente constitutive ale unui portofoliu, împreună cu obiectivele pentru care aceste activități sunt indicate. În acest fel, diferențele de la o școală la alta și de la un profesor la altul în ceea ce privește metodele de evaluare (inclusiv modul de notare) ar fi mai mici. Acest fapt nu limitează rolul profesorului, el rămânând principalul factor de decizie în stabilirea acțiunilor evaluative, succesiunii acestora și modurilor de realizare - individuale, de grup sau frontale - în funcție de particularități locale.

Evaluarea curentă utilizează, în majoritatea cazurilor, testele elaborate de profesori (vezi capitolul al IV-lea). Aceștia au și posibilitatea ca, din timp în timp, să utilizeze teste (externe) standardizate pentru a constata la ce nivel de performanță se află elevii lor față de o normă determinată în prealabil (nivel de performanță național). Această acțiune implică, desigur, existența instrumentelor standardizate de evaluare (sau cel puțin utilizarea testelor elaborate, la cerere, de Serviciul Național de Evaluare și Examinare - SNEE) și accesul ușor al fiecărei școli la obținerea lor.

În cele din urmă, modul de comunicare a rezultatelor obținute în urma aprecierilor curente influențează modul de învățare și comportamentul școlar. O notă - fie ea mare - are un efect mai redus în comparație cu o scurtă apreciere verbală sau scrisă. De aceea trebuie utilizate cât mai frecvent forme mai nuanțate de prezentare a rezultatelor, ceea ce s-a și realizat pentru ciclul primar.

6.1.1 Evaluarea în învățământul primar

Începând cu anul școlar 1998 / 1999 a fost introdusă o reformă multidimensională și profundă în întregul învățământ românesc, mai ales în ceea ce privește curriculum-ul și evaluarea. Schematic, măsurile de reformă au vizat:

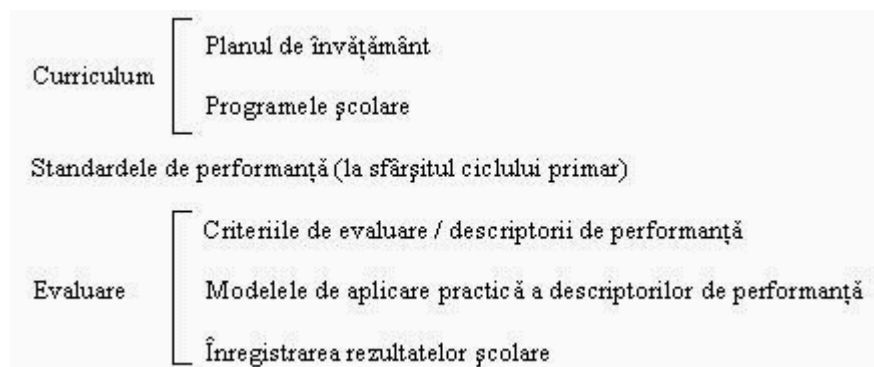


Figura 1 - Reforma curriculum-ului și evaluării în învățământul primar

Reforma evaluării elevilor din învățământul primar este o reformă de substanță, care are ca scop schimbarea completă a sistemului existent și nu simpla revizuire sau îmbunătățire a acestuia. Astfel, pentru prima dată în țara noastră s-a trecut de la evaluarea cantitativă - realizată aproape în exclusivitate în funcție de experiența cadrelor didactice și de percepția lor asupra nivelului de pregătire al elevilor - la evaluarea calitativă. Noul sistem are la bază criterii unitare, la nivel național, de apreciere a performanțelor elevilor - numite *descriptorii de performanță*. Acestea sporesc considerabil obiectivitatea evaluării, dar mai cu seamă, furnizează elevilor și părinților informații relevante despre nivelul de pregătire la care au ajuns.

Necesitatea înlocuirii sistemului de notare cifrică prin sistemul de evaluare cu calificative, bazate pe descriptorii de performanță, decurge din următoarele considerente:

a) Notele nu sunt acordate elevilor pe baza unor criterii, ci în strânsă legătură cu experiența și percepția fiecărui cadru didactic. Din această cauză, notele reprezintă simple simboluri, fără să aibă o bază științifică bine conturată.

b) În cadrul actualului sistem de notare pe o scară de zece trepte sunt imposibil de elaborat criterii relevante și unitare pentru acordarea fiecărei note. Sistemul de notare prin calificative (excelent, foarte bine, bine, suficient, insuficient), bazat pe descriptorii de performanță, asigură coerența, comparabilitatea în notare și o mai mare ușurință în utilizare.

c) Sistemul de notare prin calificative permite o evaluare mult mai obiectivă, pe niveluri de performanță superioară, medie și minimă. În acest sens, trebuie să facem o distincție clară între obiectivitatea aprecierii rezultatelor elevilor și puterea de discriminare (diferențiere) a acestor rezultate. Cel mai obiectiv sistem de evaluare este cel ce utilizează două trepte de apreciere: admis - respins.

d) Acest sistem de evaluare nu realizează o discriminare la fel de sensibilă ca sistemul cifric. În învățământul primar, considerăm că acest fapt nu este atât de important pentru evaluarea curentă, accentul urmând a fi pus mai mult pe dezvoltarea spiritului de echipă al elevilor, decât pe cel de competiție.

Evaluarea în învățământul primar are un caracter formativ, urmând ca *înregistrarea rezultatelor elevilor* în documentele școlare să se realizeze ca rezultat al mai multor evaluări (evaluare sumativă), în anumite momente stabilite de cadrul didactic. Astfel, nu orice rezultat al unei evaluări de mai mică întindere (de exemplu, răspunsurile la una - două scurte întrebări orale) este automat înregistrat în catalog, ci în caietul învățătorului. Calificativul consemnat în catalog și în carnetul elevului va fi rezultatul câtorva aprecieri de acest tip. Se pot înregistra direct în catalog calificativele obținute de elevi la evaluări cu un anumit grad de complexitate (probe scrise cu un număr suficient de cerințe, lucrări practice etc.).

În actualele documente școlare se înregistrează, în locul notelor, calificativele **Excelent, Foarte bine, Bine, Suficient, Insuficient** și datele acordării acestora. De exemplu: F.b./8.10.

Aceste calificative se acordă pe baza descriptorilor de performanță [Stoica (coord.), 1998].

6.1.2 Evaluarea în gimnaziu și liceu

În majoritatea sistemelor de învățământ din întreaga lume se acordă o atenție sporită asigurării obiectivității, transparenței și comparabilității evaluării rezultatelor școlare. În acest context, unul dintre elementele esențiale de reformă a evaluării l-a reprezentat introducerea de către Serviciul Național de Evaluare și Examinare, începând cu anul școlar 1999/2000, a unui sistem unitar de criterii pentru acordarea notelor școlare în gimnaziu și liceu.

Elaborarea acestui sistem este, în primul rând, necesară deoarece notele reprezintă doar simboluri ale unor judecăți de valoare asupra performanțelor elevilor în diferite momente ale instruirii.

În al doilea rând, sistemul de criterii asigură comparabilitatea acordării notelor de către profesori, în așa fel încât aceeași notă să reflecte niveluri de cunoștințe și competențe identice (sau foarte apropiate) pentru elevi diferiți, dar apreciați prin aceeași metodă de evaluare. [Stoica A., (coord.), 1999, p. 5 - 10].

În al treilea rând, sistemul va relaționa mai puternic evaluarea curentă de examenele naționale. Criteriile propuse vizează numai acele sarcini de lucru care suscită un grad crescut de subiectivitate în apreciere. Câteva dintre acestea sunt eseu, susținerea argumentată a unor opinii în contextul comunicării orale, proiectul, investigația ș.a.

După fundamentarea teoretică, a fost exemplificat modul de evaluare a sarcinilor menționate în cadrul diferitelor arii curriculare (Limbă și comunicare, Om și societate, Matematică și științe ale naturii). În acest sens, au fost elaborate matricele de evaluare/apreciere care cuprind criteriile și descriptorii de performanță pentru notele 5, 7, 9 și 10. Spațiile dintre nivelurile descrise (echivalentul notelor 6 și 8) pot fi acoperite cu ușurință utilizând judecata de apreciere prin comparație.

Elaborarea criteriilor de acordare a notelor se găsește în relație directă cu curriculum-ul și se realizează prin intermediul elementelor de competență definite la nivelul acestuia sau al standardelor de performanță (acolo unde există), așa cum se poate constata din schema următoare.

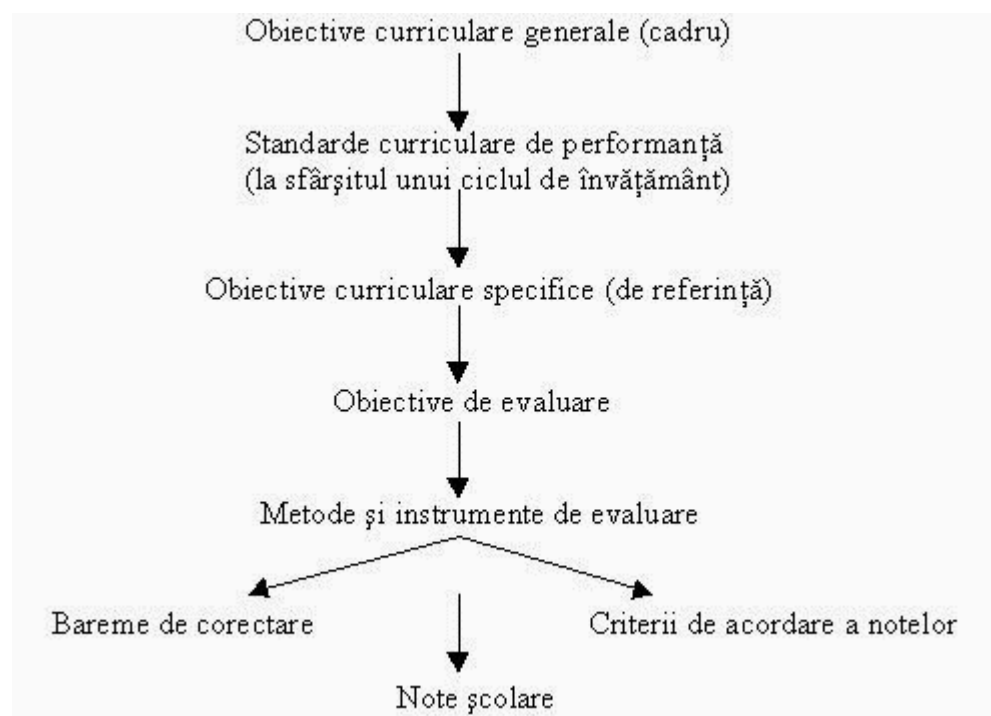


Figura 2 - Relația curriculum - evaluare

Criteriile de acordare a notelor pentru diferite sarcini de lucru reprezintă o continuare firească, într-o legătură logică și vizibilă, a sistemului de evaluare din învățământul primar.

Utilizarea matricelor de evaluare - într-o diversitate de forme și posibilități de utilizare - se poate face cu maximum de eficiență în condițiile definirii foarte clare a obiectivelor curriculare, a obiectivelor de evaluare, a sarcinilor de lucru, precum și printr-o opțiune foarte clară pentru un anumit tip de item. Potențialul major al acestor matrice de evaluare se concretizează atât prin valoarea lor formativă ridicată, în cazul elevilor care se obișnuiesc cu acest mod obiectiv și transparent de apreciere, cât și

prin stimularea și argumentarea judecății de valoare în cazul celui ce o aplică, respectiv profesorul - evaluator.

Prezentăm în continuare un exemplu. [Stoica (coord.), 1999, p. 27-30]

Disciplina: Matematică (gimnaziu)

Obiectiv de referință: Construirea unor probleme având drept suport grafice și tabele de date.

Sarcina de lucru: Pornind de la graficul și tabelele date, formulează mai multe probleme distincte, dintre care cel puțin una cu caracter practic.

Atenție! Produsul activității tale va fi apreciat după următoarele criterii:

- înțelegerea și utilizarea datelor relevante oferite;
- formularea propriu-zisă a enunțurilor;
- rezolvarea problemelor construite;
- evidențierea diferențelor (asemănări și deosebiri) între problemele create;
- argumentarea caracterului aplicativ pentru problema creată în acest scop.

Din motive de spațiu vom prezenta descriptorii de performanță (corespunzători notelor 5, 7, 9 și 10) pentru un singur criteriu și anume „formularea enunțurilor problemelor” (pentru descriptorii corespunzători celorlalte criterii se poate consulta lucrarea citată).

Criterii de acordare a notelor pentru construirea unei probleme	Nota 5	Nota 7	Nota 9	Nota 10
Formularea enunțurilor problemelor	<ul style="list-style-type: none"> • Problema este simplistă ca enunț. În construcția enunțului se regăsesc unele inversiuni, date care nu folosesc în rezolvare, formulări ambigui etc. 	<ul style="list-style-type: none"> • Enunțurile respectă structura unei probleme (ipoteză, concluzie) sunt clare, dar nu se referă la domenii diferite. Se observă tendința de a include în fiecare enunț cât mai multe din datele oferite, unele din acestea nefiind necesare în rezolvare. O problemă conține cel puțin o informație neutilizată în rezolvare. 	<ul style="list-style-type: none"> • Problemele formulate se bucură de o prezentare clară și corect realizată din punct de vedere logic. Problemele nu se referă la domenii diverse. În structura fiecărei probleme se delimitează clar ipoteza de concluzie. 	<ul style="list-style-type: none"> • Enunțurile sunt clare, precise, corecte din punct de vedere logic și al construcției frazelor. Sunt foarte bine delimitate cele două părți: ipoteza și concluzia. Enunțurile vizează domenii diferite și utilizează numai date dintre cele oferite în ipoteza de lucru. Problemele se încadrează în metode de rezolvare diferite.

6.2 Examenale: proiectare, organizare și desfășurare

În sistemele de învățământ din țările dezvoltate, dar nu numai, se dă o mult mai mare importanță (uneori absolută) finalizării unui ciclu de învățământ printr-un examen de certificare, decât admiterii într-o treaptă superioară. Nici învățământul românesc nu trebuie să facă excepție de la această abordare, având în vedere faptul că rezultatele unui examen de absolvire pot fi luate în considerare, cu o anumită pondere, pentru o viitoare selecție. Din acest motiv, Bacalaureatul și Capacitatea sunt proiectate ca examene naționale, începând cu anul școlar 1999/2000. Prezentăm mai jos câteva măsuri de reformă întreprinse care privesc aceste două examene.

Proiectarea examenului de capacitate

Fiind nou, examenul de Capacitate are avantajul că a putut fi proiectat și desfășurat de la început la un nivel calitativ superior (compatibil cu standardele internaționale), deoarece, de această dată, respectarea „tradiției”, asigurarea comparabilității subiectelor cu anii precedenți ș.a. nu au mai reprezentat obstacole în realizarea acestuia.

Câteva date le considerăm relevante pentru complexitatea, dar și pentru dificultatea proiectării și desfășurării examenului de capacitate:

Numărul de candidați: aprox. 270.000;

Numărul de școli: aprox. 13.000;

Localizarea școlilor: rural (40%), incluzând și zonele izolate;

Tipul examenului: național;

Curriculum și manualele pentru examen: programele școlare și manualele de tranziție (unice) până în anul 2000; programele școlare noi și manualele alternative începând cu examenul din anul 2001.

Disciplinele de examen: limba și literatura română și maternă, matematica, istoria românilor sau geografia României.

Principalele caracteristici ale examenului de capacitate sunt:

- a)** caracterul național al examenului, cu toate consecințele ce decurg de aici, inclusiv transformarea admiterii în liceu într-un examen local (în multe cazuri, opțional);
- b)** structura nouă a probelor de examen care să permită deplasarea accentului dinspre testarea nivelului de cunoștințe acumulate către evaluarea capacității elevilor de a utiliza aceste cunoștințe;
- c)** pilotarea/experimentarea la nivel național a examenului de capacitate;
- d)** utilizarea rezultatelor examenului de capacitate pentru admiterea în liceu;
- e)** formarea cadrelor didactice și a altor categorii de specialiști în elaborarea subiectelor de examen;
- f)** rezolvarea problemelor speciale privind organizarea examenului de capacitate (centre de examen, centre de corectare, dotări ș.a.).

Structura testelor examenului de Capacitate a fost radical schimbată față de cea utilizată pentru probele de examen desfășurate până în anul 1999. Principalele modificări au constat în: diversificarea tipurilor de itemi (pe lângă cele „tradiționale” au fost introduse întrebări cu alegere multiplă, de asociere, cu răspuns scurt, de tip eseu structurat ș.a.), realizarea unui echilibru între cerințele care solicită elevului reproducerea unor cunoștințe și cele care necesită aplicarea acestor cunoștințe, noul format al itemilor ș.a. De asemenea, baremele de corectare și notare sunt detaliate, iar notarea se realizează cu puncte de la 10 la 100. Acest fapt îl considerăm de un real folos profesorilor examinatori, care nu vor mai opera cu cunoscutele sutimi de punct.

Relația examen de capacitate - examen de admitere în liceu

Având în vedere că absolvenții clasei a VIII-a trebuiau să susțină două examene naționale (Capacitatea și Admiterea în liceu) la un interval de timp foarte scurt, la aceleași discipline și pe baza acelorași programe de examen, ponderea cea mai importantă o are examenul de Capacitate. În alte sisteme de învățământ, accentul este pus, firesc, pe „ieșirile” din sistem și mai puțin pe „intrări”. Cu alte cuvinte, învățământul românesc trebuie să diminueze treptat importanța și ponderea examenului de admitere în liceu.

Începând cu anul școlar 1998/ 1999 admiterea în liceu a devenit un examen local. Opțiunea pentru organizarea acestui examen - desfășurat pe baza unui regulament al Ministerului Educației Naționale (MEN) - revine fiecărui liceu în funcție de o serie de factori: numărul de candidați înscriși, profilul liceului, „personalitatea” liceului ș.a.

Liceele și școlile profesionale au astfel o autonomie aproape totală în realizarea examenului, în vederea selecționării elevilor cu aptitudini pentru specificul de formare al unității de învățământ respective. Schematic, situația decurge în modul următor:

Varianta 1 Liceele la care numărul de candidați înscriși este mai mic decât numărul de locuri nu organizează examen de admitere. Aceasta se realizează exclusiv în baza mediilor obținute la examenul de Capacitate.

Varianta 2 Liceele la care numărul de candidați înscriși depășește numărul de locuri organizează examenul de admitere (pentru unul sau mai multe profile) parcurg următoarele etape:

- comunică MEN / SNEE decizia de a organiza examenul și disciplina/disciplinele la care se susține proba de examen;

- SNEE elaborează teste pentru disciplina/disciplinele anunțată/anunțate;
- Liceul organizează examenul.

Nota minimă de promovare a testului de examen este 5 (cinci). Media de admitere se va calcula ca medie aritmetică a mediilor obținute de elevi la Capacitate și testul/testele de admitere. În situația în care în urma examenului de admitere există elevi respinși, situația lor va fi rezolvată prin redistribuire.

În anul școlar 1999/2000, MEN a luat decizia ca admiterea în liceu să se realizeze pe baza mediei capacității (75%) și a mediei din clasele V-VIII (25%).

Elaborarea probelor

Prima etapă în realizarea testelor pentru examenul de Capacitate a constituit-o elaborarea *programelor de examen*, printr-o atentă revizuire a programelor școlare. Capacitatea reprezintă examenul care evaluează cunoștințele și capacitățile de bază ale elevilor la sfârșitul învățământului obligatoriu. În consecință, programele de examen trebuie să cuprindă numai ceea ce este *esențial* și *util* viitorului absolvent din materia studiată.

A doua etapă constă în determinarea *structurii testelor*.

Pornind de la constatarea că elevii trebuie să demonstreze, în primul rând, capacitatea de a utiliza cunoștințele acumulate în strategii de rezolvare de probleme și abia în al doilea rând cantitatea de cunoștințe, testele trebuie să reflecte acest fapt printr-o gamă mai largă de tipuri de întrebări. Astfel, la limba și literatura română, istorie și la geografie se pot utiliza pe lângă întrebări de sinteză (de tip eseu) și întrebări „cu răspuns scurt”, „structurate” sau chiar de „tip grilă”. În acest scop, testul va conține două părți:

a) Partea 1 - întrebări structurate, cu răspuns scurt și de tip grilă;

b) Partea 2-a - întrebări cu răspuns liber (tradiționale).

A treia etapă a reprezentat-o *pilotarea națională* a examenului de Capacitate în luna februarie 1999. Serviciul Național de Evaluare și Examinare (SNEE) a organizat trei seminarii de formare (cu participarea unor experți internaționali) în lunile septembrie, noiembrie 1998 și ianuarie 1999, în vederea perfecționării cât mai multor cadre didactice, de diverse discipline, în conceperea subiectelor de examen.

Organizarea și administrarea examenului de Capacitate

Deoarece examenul de Capacitate este un examen național, organizarea și desfășurarea lui trebuie să se realizeze în aceleași condiții pentru toți candidații. O problemă specială o ridică locul de desfășurare a examenului. Este aproape imposibil din punct de vedere administrativ, logistic și financiar ca acesta să se organizeze în toate școlile generale din țară, deoarece ar fi necesare aproximativ 13.000 comisii de examen. Ca alternativă, s-a decis:

- *în mediul urban*, gruparea elevilor din mai multe școli într-o singură școală;
- *în mediul rural*, organizarea examenului la nivelul comunelor. În acest caz, o atenție deosebită a fost acordată problemei transportului elevilor, fiind necesară alocarea fondurilor și încheierea unui protocol ferm cu regiile de transport.

Utilizarea rezultatelor examenului de Bacalaureat

Examenul de Bacalaureat, ca examen național, a fost proiectat, organizat și realizat conform aceluiași etape specificate mai sus pentru examenul de Capacitate.

Pentru a stabili o strategie de utilizare a rezultatelor unei evaluări în cadrul altei evaluări trebuie să pornim de la scopul acestora.

- *Varianta întâi*

Dacă rezultatele obținute de elevi la Bacalaureat (*mediile generale*) vor fi luate în totalitate în considerare pentru admiterea în învățământul superior, atunci Bacalaureatul trebuie să-și consolideze caracterul de *examen național*, adică performanțele școlare trebuie măsurate în aceleași condiții și cu aceleași instrumente de evaluare pentru *toți elevii*. Aceasta implică:

a) elaborarea standardelor educaționale la nivel național pentru învățământul liceal, pe baza cărora se vor defini criteriile de evaluare;

- b) elaborarea unor probe de evaluare *unice* la nivel național;
- c) administrarea acestor probe în *condiții identice* (standardizate) pentru toți elevii.

Pentru a respecta aceste condiții, Bacalaureatul ar trebui să conțină numai *teste scrise*. Pe de altă parte, concentrarea Bacalaureatului numai pe teste scrise (condiție esențială pentru a-i legitima caracterul național) ar limita sever aria cunoștințelor și capacităților evaluate și ar reduce validitatea acestor probe în raport cu curriculum-ul studiat.

O altă problemă esențială este *asigurarea comparabilității* rezultatelor școlare de la un an la altul. În caz contrar, media obținută la Bacalaureat nu va avea nici o relevanță în media examenului de admitere în învățământul superior în cazul candidaților care au absolvit liceul în ani diferiți.

Așa cum am menționat mai sus, cele două examene au scopuri, moduri de proiectare și de administrare diferite. Calcularea unei medii ponderate (de exemplu, 30% Bacalaureat și 70% admitere în învățământul superior) nu-și găsește un suport științific. Facultățile ar fi, probabil, mai interesate de rezultatele obținute de elevi la Bacalaureat la disciplinele de profil specifice, decât de rezultatele globale.

În concluzie, considerăm că dacă se adoptă această variantă atunci media obținută de un elev la Bacalaureat ar trebui să constituie doar o *condiție restrictivă de înscriere la o anumită facultate*. De exemplu, Facultatea de Medicină ar urma să înscrie la examenul de admitere doar absolvenții care au obținut cel puțin media 8,50 la Bacalaureat.

• *Varianta a doua*

Facultățile ar putea lua în considerare notele obținute de elevi la Bacalaureat numai la disciplinele specialităților lor. Într-o primă etapă, aceste note ar reprezenta doar condiții restrictive de înscriere. De exemplu, Facultatea de Drept ar înscrie la examenul de admitere doar candidații care au obținut minimum 8 la probele scrise la limba română și istorie.

Intr-o etapă ulterioară, aceste note pot avea o anumită pondere în media examenului de admitere. *Scenariul* ar putea fi următorul:

a) facultățile anunță care sunt *disciplinele* la care se vor lua în considerare notele obținute la examenul de Bacalaureat și care este *ponderea* acestor note;

b) *testele scrise* pentru aceste discipline de Bacalaureat vor consta în două secțiuni:

- un test de „cunoștințe”, care să satisfacă scopul de certificare al examenului de Bacalaureat;
- un test de aptitudini generale și specifice cu caracter predictiv.

În diferite ponderi, notele obținute de elevi la cele două teste (părți ale probei) pot fi însumate fie cu media examenului de admitere, fie cu notele de la disciplinele corespondente.

Varianta a doua este, desigur, preferabilă din punct de vedere științific. Ea reprezintă o adaptare a strategiei de utilizare a rezultatelor unui tip de evaluare în alt tip de evaluare, pe plan internațional.

Pentru punerea în aplicare a acestei variante este necesară o foarte riguroasă pregătire a Bacalaureatului din punct de vedere științific și managerial. O astfel de pregătire, dacă sunt respectate toate condițiile necesare realizării unui examen calitativ și credibil, durează aproximativ doi ani. Până atunci trebuie parcursă, în mod necesar, o etapă importantă care să cuprindă:

- a) elaborarea în cadrul curriculum-ului a standardelor naționale pentru învățământul liceal;
- b) asigurarea calității științifice și pregătirea unui corp cât mai larg de specialiști în domeniul evaluării.

În plus, odată luate decizii în ceea ce privește conținutul și modul de desfășurare a Bacalaureatului, acestea trebuie mediatizate periodic și prin toate mijloacele pentru ca toți cei interesați să fie informați din timp asupra schimbărilor preconizate.

Sumar de idei

Concepte-cheie:

- evaluare curentă;
- evaluare externă;
- examene;

- testări naționale;
- descriptori de performanță;
- criterii de evaluare;
- evaluarea prin calificative;
- examen de Capacitate;
- examen de Bacalaureat;
- examene de admitere;
- pilotare;
- proiectarea, organizarea și desfășurarea examenelor;
- comparabilitatea rezultatelor școlare;
- ponderea mediilor.

Idei principale:

- Reforma evaluării rezultatelor școlare s-a materializat în trei domenii importante: evaluarea curentă, examenele naționale și evaluările naționale.
- Evaluarea curentă este parte integrantă a procesului de instruire.
- *Feedback-ul* furnizat de evaluarea curentă este util, în egală măsură, profesorului și elevului.
- Evaluarea curentă utilizează testele elaborate de profesor, dar este necesară și administrarea periodică și a unor teste standardizate.
- Evaluarea în învățământul primar a schimbat radical concepția privind aprecierea curentă a elevilor, prin trecerea de la evaluarea cantitativă la cea calitativă bazată pe descriptori de performanță.
- Descriptorii de performanță reprezintă criterii unitare, la nivel național, de apreciere a performanțelor elevilor.
- Criteriile de evaluare pentru gimnaziu și liceu au fost elaborate pentru acele sarcini de lucru care suscită un grad crescut de subiectivitate.
- Proiectarea, organizarea și desfășurarea examenelor naționale implică parcurgerea unor etape clare, standardizate, pentru ca aceste examene să devină credibile.
- Examenele naționale au fost proiectate în ultimii ani de către Serviciul Național de Evaluare și Examinare, instituție specializată în evaluarea educațională (aflată în subordinea MEN)
- Rezultatele examenelor de Capacitate și Bacalaureat constituie baza admiterii în treptele superioare de învățământ (liceu, universități).
- Transparența organizării și desfășurării examenelor, informarea permanentă a celor interesați privind modificările preconizate și formarea cadrelor didactice în domeniul evaluării sunt condiții esențiale pentru punerea în practică a unui sistem credibil și obiectiv de evaluare curentă și examinare.

GLOSAR MINIMAL DE TERMENI

Argumentul pentru scrierea acestui glosar minimal de termeni îl reprezintă dorința autorilor de a evita neînțelegerile de sens provocate de utilizarea diferitelor concepte operaționale în acest ghid. Având în vedere faptul că termenii pot avea diferite accepțiuni pentru diferiți specialiști în funcție de experiența educațională a acestora sau de context, **glosarul** propune definițiile de lucru care au fost utilizate în conținutul prezentului ghid.

TERMEN	DEFINIȚIA DE LUCRU
Banca de itemi	o colecție de itemi organizați după o anumită structură și clasificați după criterii clar stabilite, din care aceștia pot fi aleși cu scopul de a elabora teste. Itemii din componența băncilor de itemi trebuie mai întâi experimentați.
Criteriu	normă conform căreia o probă poate fi apreciată sau evaluată; descrierea unei performanțe.
Efect de backwash	influența exercitată de examene asupra curriculum-ului.
Efect de feedback	circulația permanentă a informației obținută printr-o anumită tehnică de evaluare către profesori, factori de decizie, părinți etc.
Eșantion reprezentativ	eșantion care corespunde populației din care este extras în ceea ce privește caracteristicile importante ale acesteia, oferind posibilitatea generalizării concluziilor asupra întregii populații.
Evaluare națională	evaluare care determină în ce măsură nivelul de performanță al elevilor la sfârșit de ciclu sau perioadă importantă de instruire corespunde standardelor educaționale la nivel național, specifice ciclului de învățământ respectiv.
Fidelitate	calitatea unui instrument de evaluare de a produce rezultate constante în urma aplicării sale repetate.
Instrument de evaluare	parte integrantă a metodei de evaluare, care concretizează la nivel de produs opțiunea metodologică a profesorului pentru testarea performanțelor elevilor într-o situație educațională bine definită.
Item	întrebare și răspunsul așteptat la aceasta.
Matrice de specificații	o matrice care specifică formatul unui instrument de evaluare.
Metoda de evaluare	ansamblul operațiilor/etapelor de urmat pentru atingerea obiectivului/obiectivelor de evaluare propus(e), utilizând un instrument de evaluare adecvat.
Nivel de performanță	gradul de reușită individuală.
Notare analitică	tip de notare aplicată în general testelor de tip formativ, presupunând acordarea de puncte/scoreuri separate pentru fiecare dintre elementele importante ale răspunsului (criterii specificate, ca de exemplu: corectitudinea punctuației, fluența scrisului etc.).
Notare holistică	tip de notare care vizează aprecierea răspunsului în totalitatea sa și încadrarea acestui răspuns într-o categorie prestabilită.

Obiective de evaluare	tip de obiective care sunt echivalente obiectivelor operaționale, adică trebuie să evidențieze cele trei elemente ale operaționalizării: comportamentul observabil, condițiile în care se produce comportamentul respectiv, precum și nivelul de performanță sau criteriul de reușită. Prin intermediul lor, rezultatele învățării sunt exprimate în termeni direct măsurabili.
Obiectivitatea	calitate a unui test care se referă la gradul de concordanță între aprecierile făcute de către evaluatori independenți în ceea ce privește un răspuns bun pentru fiecare dintre itemii testului.
Portofoliu	instrument de evaluare complex și integrator care include rezultatele relevante obținute prin celelalte metode și tehnici de evaluare (probe orale, scrise, practice, proiect, autoevaluare etc.)
Probă	orice instrument de evaluare proiectat, administrat și corectat de către profesor.
Programă de examen	document care specifică ceea ce trebuie evaluat în situație de examen și cum anume.
Scorul real	Scorurul/punctajul obținut în urma aplicării unui instrument de evaluare.
Scor observat	scor real \pm eroarea scorului.
Test criterial	tip de test pentru care aprecierea rezultatelor elevilor se face în raport cu un criteriu/o serie de criterii de performanță stabilit(e) anterior.
Test normativ	tip de test pentru care rezultatele înregistrate sunt evaluate în raport cu cele ale unui grup de referință, nu cu un criteriu stabilit anterior.
Test standardizat	tip de test care cuprinde itemi obiectivi și semiobiectivi care sunt selectați dintr-un număr foarte mare de itemi, grupați în funcție de diisciplină, nivel de vârstă, obiective, nivel de dificultate etc. în bănci de itemi.
Validitate	calitatea unui instrument de evaluare de a măsura exact ceea ce și-a propus să măsoare.

BIBLIOGRAFIE

1. Ausubel, D.; Robinson, F., (1981) Învățarea în școală, București: Editura Didactică și Pedagogică.
2. Campbell, Jay R.; Voelkl, Kristin E.; Donahue, Patricia L. (1997) Report in brief. NAEP 1996.
3. Carey, L. (1988) Measuring and Evaluating School Learning. Boston: Allyn and Bacon, Inc.
4. Carey, L. (1988) Measuring and Evaluating School Learning. Allyn and Bacon, Inc.
5. De Landsheere, Viviane; De Landsheere, Gilbert (1997) Definirea obiectivelor educației. Tr. de Constantin Urma. București: Editura Didactică și Pedagogică.
6. Fairbrother, R.V., Watson, J.R., Jones, A.T. (1992) Open Work in Science. Inset for investigations. University of London.
7. Frith, D.S.; Macintosh, H.G. (1984) A Teacher's Guide to Assessment. Glasgow: Bell a Bain. Stanley Thornes Publishers
8. Frith, D.S.; Macintosh, H.G. (1991) A Teacher's Guide To Assessment. Londra: Stanley Thornes, Ltd.
9. Gipps, Caroline (1994) Beyond Testing: Towards a Theory of Educational Assessment. London: The Falmer Press
10. Gronlund, Norman, E. (1981). Measurement and Evaluation in Teaching. Ed. a IV-a. N.Y.: MacMillan Publishers Co.inc.
11. Gronlund, N. (1998) Measurement and Evaluation in Teaching (sixth edition). Boston: Allyn and Bacon
12. Haladyna, T.M. (1997) Writing Test Items to Evaluate Higher Order Thinking. Boston: Allyn and Bacon
13. Jacobs, L.C; Chase, C.I. (1992) Developing and Using Test Effectively. A Guide for Faculty. San Francisco: Jessey-Bass Publishers
14. Krasnaseschi, Vi. (coord.) (1976) Evaluarea sistemelor și a proceselor educaționale. București: E.D.P.
15. Neacșu, I., Stoica, A. (coord.) (1996) Ghid general de evaluare și examinare. București: Ed. Aramis
16. Osterlind, S.J. (1998) Constructing Test Items: Multiple Choice, Constructed-Response, Performance, and Other Formats. Boston: Kluwer Academic Publishers
17. Patton, Michael Quinn (1987) Creative Evaluation. Ed. a II-a. London: SAGE publications, The Publishers of Professional Social Science.
18. Potolea, D., Neacșu, I., Radu, I.T., (1996) Reforma evaluării în învățământ. Concepții și strategii. București, Consiliul Național de Evaluare și Examinare
19. Radu, I.T. (1981). Teorie și practică în evaluarea eficienței învățământului. București: E.D.P.
20. Roid, G.H.; Haladyna, T.M. (1982) A Technology for Test – Item Writing. Orlando: Florida, Academic Press
21. Somerset, H.C.A., (1993) Examination and the Quality of Education. The World Bank (Interim Report)
22. S.N.E.E. (2000) Modele de teste
23. Tenmark, J. (editor) (1991) Mathematics Assessment. Virginia: The National Council of Teachers of Mathematics Inc.
24. Stoica, A. (coord.) (1998) Evaluarea în învățământul primar. Descriptori de performanță. București: Editura Humanitas.
25. Stoica, A., Musteață, S. (1997) Evaluarea rezultatelor școlare. Chișinău: Ed. Liceum.
26. Stoica, A. (2000) Reforma evaluării în învățământ. București: Ed. Sigma.
27. Stoica, A. și alții (1999) Un sistem de criterii pentru acordarea notelor în învățământul gimnazial și liceal. București: Editura Școala Românească.
28. Thyne, J.M., (1978) Examinarea elevilor (traducere). București: Editura Didactică și Pedagogică.
29. Trends in Academic Progress. U.S. Washington, D.C.: National Center for Educational Statistics.
30. Walberg, Herbert J.; Haertel, Geneva D. (eds.). (1990) The International Encyclopedia of Educational Evaluation. Oxford, England: Pergamon Press Ltd.
31. XXX (1998) Curriculum Național pentru învățământul obligatoriu. Cadru de referință. București, Consiliul Național pentru Curriculum – M.E.N.
32. XXX (1987) Glossaire de termes relatifs à l'évaluation et termes connexes. Paris: UNESCO.

CUPRINS

Prefață	3
Capitolul I: Evaluarea educațională și evaluarea performanțelor / rezultatelor școlare	4
Capitolul II: Relațiile dintre curriculum și evaluare	22
Capitolul III: Metode și instrumente de evaluare	30
Capitolul IV: Elaborarea și administrarea testelor scrise	47
Capitolul V: Tipologia itemilor	65
Capitolul VI: Reforma evaluării rezultatelor școlare în România	86
Glosar	94
Bibliografie	96